

UNCLASSIFIED



Australian Government
Department of Defence
Defence Science and
Technology Organisation

Moving Forward with Computational Red Teaming

Phillip Gowlett

Joint Operations Division
Defence Science and Technology Organisation

DSTO-GD-0630

ABSTRACT

This paper assesses the requirements for developing Computational Red Teaming for application to Counter Improvised Explosive Devices and other defence problems. Key activities contributing to this report include the delivery of four research papers from Australian Universities engaged through research agreements, and a gathering of the broader community of interest at a workshop held at the Joint Decision Support Centre. The research was conducted under the Counter Improvised Explosive Device Corporate Enabling Research Program of FY2009.

RELEASE LIMITATION

Approved for public release

UNCLASSIFIED

UNCLASSIFIED

Published by

*Joint Operations Division
DSTO Defence Science and Technology Organisation
Fairbairn Business Park Department of Defence
Canberra ACT 2600 Australia*

Telephone: (02) 6265 9111

Fax: (02) 6128 6332

© Commonwealth of Australia 2011

AR-014-953

March 2011

APPROVED FOR PUBLIC RELEASE

UNCLASSIFIED

Moving Forward with Computational Red Teaming

Executive Summary

Joint Operations Division (JOD) is investigating the emerging concept of Computational Red Teaming (CRT) for the purpose of enhancing decision support and red teaming (RT). Increasing compute power and agent based architectures that allow for the creation and exploitation of intelligent agents capable of learning, create the potential for more complete explorations of complex issues and scenarios.

In order to understand the merits and challenges of different approaches to CRT, JOD has engaged a research community of interest, and contracted universities to scope the topic area by way of literature and computational tool reviews. These reports were presented at a workshop where both defence researchers and academics tabled their findings, thoughts and interests.

The workshop indicated a need for research and development in the following areas in order to apply CRT.

- Social models governing blue-white-red interactions in a counter-insurgency context.
- Cognitive models governing red-white-blue planning and actions.
- Search algorithms and methodologies to determine the impact of varying key parameters on a large and complicated problem domain or scenario space.

JOD will need to develop expertise in these areas in order to construct the kind of CRT capability researchers envisage for the Counter Improvised Explosive Device Domain.

Other information that emerged from the workshop includes:

- a prototype CRT capability under development for CIED in the Defence Operations Support Centre;
- a possible role for purpose built games and computation through play to support military planning.

Experience in defence RT will be important for successfully designing and standing up a CRT capability. For CRT to successfully influence the decision making process, client interest and backing of RT is critical. Through discussions with JOD analysts involved in military experimentation, it has been conveyed that RT is an established part of tactical and operational ADF planning and experimentation but the benefits are not being fully realised at the strategic level.

UNCLASSIFIED

Recommendations for CRT Research and Development

- Demonstrate a proof of principle CRT capability in a simple problem domain.
- Facilitate JOD CRT researcher involvement in client endorsed RT to better understand the practice, practical issues, utility and limitations of RT in the ADF.
- Support the generation of expertise in and advancement of thick agents, evolutionary algorithms and co-evolution models.
- Investigate how thick agents and co-evolution can be interfaced with existing physical models and identify the challenges posed.
- Explore the potential of CRT to contribute to Operations Analysis and Operations Research.
- Identify a champion willing to support the implementation of CRT approaches in the client space.
- Investigate the potential utility of purpose built games and computation through play approaches, possible implementation schemes, their feasibility and likely resource requirements.
- Investigate the relevance of adversarial learning to CRT.

Recommendations for the application of CRT to CIED

- Consider the need to bring together social scientists and operational data for the generation of cognitive and social models for application to CIED and counter-insurgency.
- Support the generation of expertise in and advancement of cognitive and social modelling
- Support the CIED threat anticipation modelling efforts underway in the Defence Operations Support Centre; explore its potential role as a test bed or development platform for thick agents, evolutionary algorithms and coevolution.

Within JOD, there are a number of disciplines that could benefit from agent based modelling techniques. Further recommendations are made within the report to support the development of JOD expertise in these methods.

UNCLASSIFIED

UNCLASSIFIED

Author

Phillip Gowlett

Joint Operations Division

Phillip Gowlett joined DSTO in 2007 having been awarded the qualifications of Bachelor of Aerospace Engineering and a Bachelor of Science with Honours in Physics from the University of New South Wales in 2006. Within DSTO, his experience lies in the conduct of studies and analysis falling within the needs phase of the capability development cycle for clients in the Capability Development Group.

UNCLASSIFIED

Contents

ACRONYMS

1. INTRODUCTION	1
1.1 Background.....	1
1.2 Purpose	2
2. LITERATURE	3
2.1 Introductory Concepts	3
2.1.1 Adaptable algorithms	3
2.1.2 Agent based distillations & models.....	3
2.1.3 Co-evolution	3
2.1.4 Fitness landscapes.....	3
2.1.5 Game theory	4
2.1.6 Nash Equilibrium	4
2.1.7 Serious games and computation through play	4
2.1.8 Thick agents	4
2.2 Defining CRT.....	4
2.3 Agent Based Modelling.....	5
2.4 Academic Papers	6
3. WORKSHOP	7
3.1 Prof Hussein Abbass & Dr Michael Barlow, ADFA.....	8
3.2 Prof Peter Campbell & Dr Mathew Berryman, UniSA	8
3.3 A/Prof Philip Hingston, ECU	8
3.4 A/Prof Cara MacNish, UWA.....	9
3.5 Dr Axel Bender, LOD.....	9
3.6 Dr Tony Dekker, JOD.....	9
3.7 Mr Justin Millikan, LOD.....	10
3.8 Mr Nathan Sayers, JOD	10
4. DISCUSSION.....	10
4.1 Consensus.....	10
4.2 Challenges	11
4.3 Blind spots	13
5. RECOMMENDATIONS.....	14
5.1 General CRT Research and Development	14
5.2 CIED application of CRT	14
5.3 Joint Decision Support Centre / JOD specific.....	15
APPENDIX A: WORKSHOP ATTENDANCE	16
APPENDIX B: ACADEMIC PAPERS	17

UNCLASSIFIED

ACRONYMS

ABD	Agent Based Distillation
ABM	Agent Based Model
ADF	Australian Defence Force
ADFA	Australian Defence Force Academy
ART	Auto Red Teaming
CDG	Capability Development Group
CIED	Counter Improvised Explosive Device
CRT	Computational Red Teaming
DSTO	Defence Science and Technology Organisation
ECU	Edith Cowan University
IED	Improvised Explosive Device
JDSC	Joint Decision Support Centre
JOD	Joint Operations Division
LOD	Land Operations Division
OA	Operations Analysis
OR	Operations Research
RT	Red Teaming
UniSA	University of South Australia
UWA	University of Western Australia

UNCLASSIFIED

1. Introduction

Red teams and red teaming (RT) processes have been used as tools by both government and commercial organisations to identify and reduce risks. RT is used to challenge aspects of the enterprise's plans, programs and to test assumptions. It is this aspect of deliberate challenge that distinguishes RT from other tools. RT can help to hedge against surprise, particularly surprises with catastrophic consequences. Traditionally RT has been a human centric activity that relies upon creativity, intuition and subject matter expertise in order to put a set of assumptions or decisions to the test.¹

Software advances and computational tools have the potential to enhance RT and its value to decision support. As problems become more complex humans struggle to keep track of the complete set of variables, parameters and factors that are intrinsic to the problem. Additionally, the size of the options set to be investigated becomes unmanageable, due to time and resource constraints. Computational techniques have the potential to address these issues via modelling and simulation combined with numerical search and optimisation. A broad array of linked topics have been identified during a preliminary scoping phase including artificial markets, statistical learning, data farming, co-evolution, risk assessment, game theory, games with improvised explosive devices (IED) scenarios, agent based distillation (ABD), fitness landscape and adversarial learning.

Computational Red Teaming (CRT) is a new concept yet to be implemented. There is no formal consensus on a definition of CRT. A key aim of CRT is to enhance the quality of decision making by increasing the degree of rigor that can be brought to bear on complex problems. Three possible avenues for CRT to do this include:

- Supporting RT through the generation of knowledge for use by red teams in experimentation.
- Use of software tools to help structure the RT process and provide greater information to participants.
- Conducting the RT process by semi-automated computational methods.

1.1 Background

The CRT Corporate Enabling Research Program (CERP) activity aims over FY 2009–10 are to:

- Create a research and applications community of interest jointly shared across DSTO and academia, emphasising Defence applications of CRT in decision support.
- Scope computational approaches to assist with the prediction phase of CIED and the impact of new blue capability on red behaviours.
- Develop a research program for FY2010–11 targeting promising areas identified during the scoping phase.

¹ *The Role and Status of DoD Red Teaming Activities*, Office of the Undersecretary of Defence, 9/2003, <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA430100> accessed 10/2010

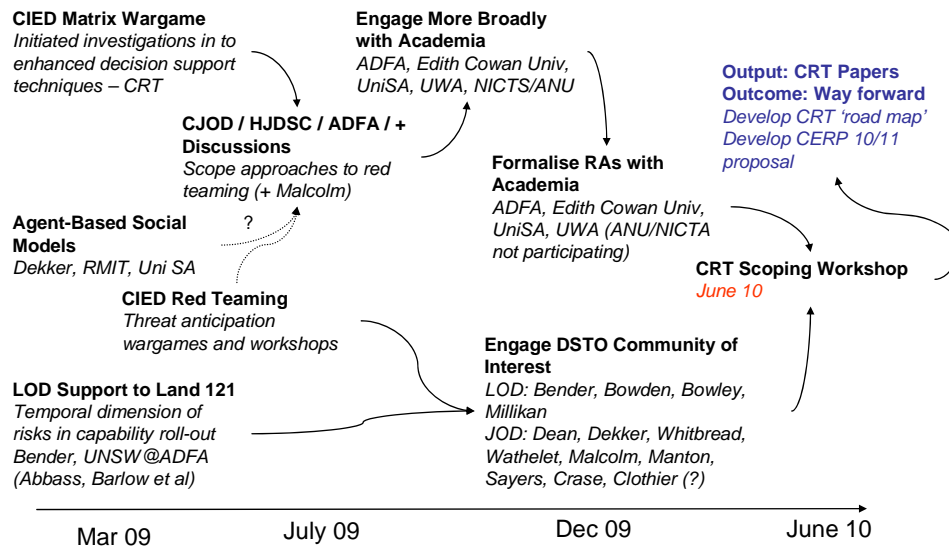


Figure 1: Timeline of CRT related activities over 2009–10

Late in 2009, formal research agreements were signed with the Australian Defence Force Academy (ADFA), The University of South Australia (UniSA), the University of Western Australia (UWA) and Edith Cowan University (ECU). The agreements were for literature and computational tool reviews to be delivered to the Defence Science and Technology Organisation (DSTO) in formal reports, and presented at a DSTO workshop. The context given was CRT as applied to the CIED domain.

1.2 Purpose

This paper evaluates the findings of the four university literature and computational tool reviews and discusses emerging themes from the greater body of literature. Issues raised in the subsequent workshop that brought together the various CRT interest groups for the first time were documented and examined. The information provided by these activities was evaluated and recommendations provided concerning the future directions of CRT research activities within JOD.

2. Literature

2.1 Introductory Concepts

A number of technical terms, techniques and research fields are referred to throughout this paper. For the benefit of the reader their definitions within the context of this paper are outlined below.

2.1.1 Adaptable algorithms

The term adaptable algorithm essentially refers to biased random search methods. These search methods are useful when dealing with fitness landscapes that are too large to calculate the entire solution set and too rough to apply simple hill climbing search approaches. Three different search techniques falling within this class are:

- Particle swarm optimisation
- Evolutionary algorithms
- Simulated Annealing

2.1.2 Agent based distillations & models

ABDs are simulation approaches that model warfare through the interactions of agents in a scenario space. Distillations are much less detailed than traditional simulations. Distillations implement a small rule set and generate insights through the emergence of global behaviour from the interactions of many agents. Distillations allow a far closer modelling of complex adaptive systems than any other simulation paradigm. Distillations are a bottom up approach to modelling complex systems. Agent based distillations are distinguished from agent based models and simulations by their minimalist level of fidelity. The term agent based models or simulations might also be used to refer to the field of agent based modelling generally.

2.1.3 Co-evolution

Co-evolution is the process where two related species or entities adapt to each others adaptations. In this context one can co-evolve two military forces against each other within an agent based model using biased random search methods to optimise an agent's parameters and behaviour. Each force will adapt in such a way as to defeat the previously encountered adversary.

2.1.4 Fitness landscapes

Fitness landscapes are plots of all possible solutions to a given problem. It is important to note that the number of solutions for the kinds of military problems investigated with agent based models is extremely large, such that the task of calculating the entire solution set may not be possible. Finding the best solution in a large fitness landscape is a challenging and relevant problem for CRT that is likely to require the application of biased random search techniques.

2.1.5 Game theory

Game theory is an applied branch of mathematics that attempts to mathematically capture behaviour in situations where an individual's success in making choices depends on the choices of others. Game theory is used to investigate how one should make decisions in these situations and to a lesser extent how one does make them.²

2.1.6 Nash Equilibrium

The Nash Equilibrium is a solution concept of a game or situation involving multiple players, where no player has anything to gain by changing their own strategy unilaterally. If each player has chosen a strategy and no player can benefit by changing this strategy while the others keep theirs unchanged, then the current set of strategy choices and corresponding payoffs constitute a Nash equilibrium.

2.1.7 Serious games and computation through play

Serious games are the application of entertainment software or the design and development principles of such software for non-entertainment purposes. An example is the use of first person perspective computer game technology for military training and experimentation. Computation through play is the process whereby a research question is explored via a purpose designed game construct. Appropriate rules and scenarios are constructed in the game, and data is collected concerning in game events and analysed to derive insights relevant to the research problem. These techniques are potential avenues for tapping the wisdom of crowds through a virtual environment.³

2.1.8 Thick agents

In the context of CRT, agents are software agents that perceive their environment and then act in a goal directed manner interacting with other software in an autonomous fashion. 'Thick' refers to the level of sophistication of the software agent. A simple agent might move or take an action under certain circumstances such as its proximity to another type of agent and a simple goal which can be represented by a few lines of code. A thick agent has much greater software sophistication that enables it to perceive more information. Additionally the agent may possess a learning mechanism. This allows thick agents to determine their actions through the use of dynamic cognitive models rather than static deterministic if/then decision trees in the case of simple agents.

2.2 Defining CRT

*'CRT is a set of methodologies and computational models that augment a human based red teaming exercise or perform a computer based, more abstract red teaming exercise.'*⁴

² M Davis, *Game Theory A Nontechnical Introduction*, Dover Publications, Mineola, New York, 1983, pp. 3

³ H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, Defence and Security Applications Research Centre, University of New South Wales @ the Australian Defence Force Academy, 5/2010, pp. 5–8

⁴ H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, Defence and Security Applications Research Centre, University of New South Wales @ the Australian Defence Force Academy, 5/2010, pp. 2

*'...CRT is a framework built upon a set of computational models that can assist a human based red teaming exercise smartly and responsibly.'*⁵

The current attempts to define CRT in the Australian context as seen in the quotation above are very broad and could include practically any computational tool or technique a red team finds helpful, for example: a reference database of platform capabilities, pre-run combat engagements or a visualisation tool to assist with information transfer.

The body of published literature that specifically treats CRT is not yet extensive and much of it is authored by Prof. Hussein Abbass who has been engaged for the purpose of providing a literature and tool review for this paper. However, considerable related literature exists spanning areas such as computational decision support, applications of gaming in operations research and agent based modelling and simulation. Rather than duplicate the work of the academics the author refers the reader to Appendix B for an extensive treatment of the literature. Section 2.4 provides a brief overview of the approach and content of the literature reviews.

2.3 Agent Based Modelling

A range of agent based models have been developed and applied to investigating warfare since the 90s. ISAAC (Irreducible Semi-Autonomous Adaptive Combat) and EINStein (Enhanced ISAAC Neural Simulation Toolkit) are among the first and inspired many of the ABDs that were later developed. MANA (Map Aware Non-uniform Automata) which was developed at the New Zealand Defence Technology Agency has been used to analyse military problems in many countries. Notably MANA introduced way-points and map and event driven personality changes⁶. Other models such as BactoWars (developed in LOD, DSTO), WISDOM and Socrates have subsequently been developed to cater for specific needs and to address the limitations of existing models. NetLogo is today one of the most widely used agent based modelling architectures across a range application areas and is freely available for download. Agent based models have successfully enabled analysts to gain an understanding of what factors are most important to the outcome of conflicts. Agent based models are an efficient way to create and investigate complex systems and behaviour such as military forces and their interactions in combat.

⁵ H Abbass, A Bender, P Whitbread, *Computational Red Teaming: Unravelling the Pharaoh's Curse*, Computational Intelligence in Security and Defence Applications Workshop, WCCI2010, Barcelona, Spain

⁶ *Limitations of agent-based distillation systems*, http://epress.anu.edu.au/cs/mobile_devices/ch08s03.html accessed 11/2010

UNCLASSIFIED

A number of international papers have coined the term *Auto Red Teaming* (ART); ART is also highlighted on the Project Albert website.⁷ ART is an approach utilising agent based simulations and evolutionary algorithms to explore a scenario space in order to gain insights into optimised strategies for either one or both sides.⁸ It is clear that ART falls within the bounds of the broader definition of CRT. In ART, humans do not shoulder the cognitive load in the problem exploration stage, only in the problem definition, configuration and post exploration analysis. This interpretation or aspiration for CRT was evident in a number of the workshop presentations, as academics showed support for research and development of automated software for application to complex problem domains. Of interest to researchers is the possibility of utilising co-evolution and evolvable environments and scenarios to further expand the dynamics that can be explored with the ART modelling approach.⁹

A look across the literature demonstrates that ART research and development makes use of ABDs to model relatively simple problems containing modest numbers of actors and variables. These problems are highly abstracted and used as proof of principle test cases for the concept of ART. Examples include an Urban Operation⁸ consisting of raiding and capturing of a deliberately defended location in the presence of civilians, and a maritime anchorage defence scenario⁹ where red attack vessels must breach a blue defensive patrol and inflict damage on commercial vessels within the anchorage.

Despite initial attempts, the benefits of co-evolution and evolvable scenarios have yet to be demonstrated in a compelling manner in ART papers reviewed for this body of work. It is also not clear that evolutionary algorithms are the best optimisation technique for the kinds of military problems described above¹⁰.

2.4 Academic Papers

Below is a brief overview of the approaches and directions of the academic literature and tool reviews contracted from four universities to inform this paper on the task of assessing CRT and its application in supporting CIED. The academic papers in full can be found in Appendix B.

The paper¹¹ from UNSW@ADFA provides the deepest consideration of what CRT is, how it has come about, why it's needed and what differentiates it from other research. The paper draws a link between CRT and serious games via simulation and then explores and advocates the use of serious games and computation through play as a potential simulation approach under the CRT framework, highlighting a few examples of relevance to CIED.

⁷ Project Albert website: <http://www.projectalbert.org/index.html> accessed 10/2010

⁸ C Choo, C Chua, S Tay, *Automated Red Teaming, A Proposed Framework for Military Application*, GECCO'07, 7/2007

⁹ Y Xu, M Low, C Choo, *Enhancing Automated Red Teaming with Evolvable Simulation*, GECCO'09, 6/2009

¹⁰ C MacNish, *Computational Red Teaming: A Review for the Defence Science and Technology Organisation*, Faculty of Engineering, Computing and Mathematics, The University of Western Australia, 5/2010, pp. 9–18

¹¹ H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, Defence and Security Applications Research Centre, University of New South Wales @ the Australian Defence Force Academy, 5/2010

The UniSA paper¹² is strongly oriented towards addressing questions of relevance to the CIED problem. The paper bypasses any deep consideration of what CRT is with a short declaration of its interpretation in the introduction. The paper then identifies a key aspect of CIED (social behaviour) that if better understood could lead to better operational outcomes before relevant literature and a range of possible modelling approaches are critiqued and discussed. The paper concludes by putting forward agent based models (ABMs), thick agents, time ordered event queues and separation of model and modelling architecture as important constituents of a CRT CIED capability.

The ECU paper¹³ focuses on the ART interpretation of CRT as it might be applied to CIED with or without use of co-evolution methods. A simple (in-house) example that applies MANA to an IED attack scenario is described and the lessons from this are discussed. A general discussion of agent based modelling and simulation is provided, ranging across areas such as optimisation methods, evolutionary algorithms, data mining, and the concepts of Nash Equilibrium and fitness landscape. A range of ABMs and ABDs are then discussed in a general manner.

The UWA paper¹⁴ begins with a comparison of the strengths and weaknesses of RT and modelling and simulation approaches. Deriving insights from this comparison for the purpose of framing an interpretation of CRT, the paper then focuses on the progress towards employment of techniques from artificial intelligence, computational intelligence, adaptive and agent based systems. The paper assesses a range of literature in these areas, in the style of a narrative that details the author's journey through the literature. In conclusion, the need to combine adaptive algorithms with social cognitive modelling is seen to be of key importance for progress toward CRT CIED capabilities.

3. Workshop

DSTO and academic researchers interested in CRT were brought together at a conference style workshop at the Joint Decision Support Centre (JDSC) in Canberra on the 7th of June 2010. Researchers presented their work and perspectives on CRT and then fielded questions. A short session was provided at the end of the day for general discussion on the topic of future research efforts. What follows is a list of the key points each presenter contributed along with some comment on the risks of implementing their approach. A list of the attendees can be found in Appendix A.

¹² P Campbell, *A Literature Review for IED Computational Red Teaming*, Defence and Systems Institute, University of South Australia, 5/2010

¹³ P Hingston, *Computational Red Teaming – A Literature Survey and Computational Tool Review*, School of Computer and Security Science, Edith Cowan University, 4/2010

¹⁴ C MacNish, *Computational Red Teaming: A Review for the Defence Science and Technology Organisation*, Faculty of Engineering, Computing and Mathematics, The University of Western Australia, 5/2010

3.1 Prof Hussein Abbass & Dr Michael Barlow, ADFA

- Justification and reasoning for CRT is put forward.
- The challenges and issues associated with the conduct of CRT are considered.
- CRT is characterised in-terms of its scope, boundaries and context.
- Reservations were expressed over the readiness or suitability of evolutionary algorithms in the context of CRT for CIED.
- Advocates serious games and computation by play as a way forward.
 - *Technically feasible with mature, available technologies.*
 - *Costly to implement.*
 - *Not yet attempted at significant scale for the purpose of generating data to support analysis.*
 - *DSTO does not have the technical capabilities required and would need to partner with academia & commercial software developers.*¹⁵

3.2 Prof Peter Campbell & Dr Mathew Berryman, UniSA

- Social networking representations are recognised to be of key importance for CIED modelling.
- Modelling the multidimensional heterogeneous CIED domain is probably best done using an agent based architecture, time ordered event queue and separating what is modelled from how it is modelled.
- A move towards thicker agents and away from ABDs is recommended.
- Advocates thick agent based models and social network analysis.
 - *Capable of sophisticated behaviour.*
 - *Models are data hungry.*
 - *Thick agents, cognitive and social modelling are not presently readily accessible and applicable in JOD.*
 - *JOD's lack of social science expertise is a constraint.*¹⁶

3.3 A/Prof Philip Hingston, ECU

- CRT is recognised as a means to conduct a much more complete search of the problem space.
- Searching for the Nash Equilibrium provides the most robust solution.
- Evolutionary algorithms are suited to complex multi-objective problems requiring the generation of an option set.
- ABDs are no longer necessary due to advancements in computational power.
- Advocates the application of agent based simulations, evolutionary algorithms and conditionally co-evolution.
 - *ART approach to CRT.*
 - *Ensuring fit for purpose representation of human factors a challenge.*
 - *Limited comment on the application of this approach to CIED problems.*
 - *No sophisticated examples of this approach yet demonstrated.*¹⁷

¹⁵ H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, PowerPoint presentation to JDSC Computational Red Teaming Workshop, 6/2010

¹⁶ M Berryman, P Campbell, *Computational Red Teaming*, PowerPoint presentation to JDSC Computational Red Teaming Workshop, 6/2010

3.4 A/Prof Cara MacNish, UWA

- Human factors are important as individuals are capable of generating significant effects in counter-insurgency and CIED operations.
- Proof of concept demonstrations of ART to date fall short of complexity required to represent CIED problem.
- There is significant development underway in cognitive modelling and modelling social interactions.
- Advocates need to combine sophisticated modelling environments with cognitive and social models.
 - *ART inspired approach to CRT.*
 - *Need to understand the fitness landscape for application of the correct algorithms (may or may not be evolutionary algorithm).*
 - *More research is needed to move towards applications.*¹⁸

3.5 Dr Axel Bender, LOD

- CRT should be focused on testing the fitness of decisions.
- A range of CRT principles were identified and discussed including:
 - *Adaptability*
 - *Handles multi-objective optimisation*
 - *Reflectivity & Reflexivity (co-evolution)*
 - *Decision testing*
 - *Risk based justification*
 - *Participatory*
 - *Interdisciplinary*
- Foundational issues concerning what CRT is and isn't that in turn influence what the appropriate application areas for CRT were discussed.¹⁹

3.6 Dr Tony Dekker, JOD

- CIED domain focus.
- Advocate for the use of agent based models.
- Real need to represent human factors and social networks in CIED modelling.
- The components for a CIED CRT capability have been demonstrated separately in different research centres but not yet integrated.
- Best search methodology to apply to resultant fitness landscapes is not clear.
- Aligns with the ART inspired approach to CRT.²⁰

¹⁷ P Hingston, *Computational Red Teaming*, PowerPoint presentation to JDSC Computational Red Teaming Workshop, 6/2010

¹⁸ C MacNish, *Computational Red Teaming*, PowerPoint presentation to JDSC Computational Red Teaming Workshop, 6/2010

¹⁹ A Bender, *Framework for CRT Design*, presentation to JDSC Computational Red Teaming Workshop, 6/2010

²⁰ T Dekker, *Agent-Based Simulation for Counter-IED: A Simulation Science Survey*, presentation to SimTecT2010 & JDSC Computational Red Teaming Workshop, 6/2010

3.7 Mr Justin Millikan, LOD

- jSWAT (campaign level), JANUS (Brigade level), CASTFOREM (Battalion level) – facilitation of adjudicated seminar war gaming.
- Double blind multi-side war games.
- Bactowars – long running LOD research program to produce an ABM capability.
- CHAN – section sized forces, detailed urban terrain.
- Significant and ongoing use of agent based modelling to investigate land warfare.²¹

3.8 Mr Nathan Sayers, JOD

- A CIED threat anticipation model is under development that incorporates a social model at the insurgent cell level.
- The social model is informed by operational data rather than BDI agents²².
- Light on automation – simple reactive agents and no evolutionary algorithms which necessitates human in the loop optimisation/search.
- Focused toward application supporting military intelligence & planning in the CIED domain.²³

4. Discussion

4.1 Consensus

There does appear to be some consensus among DSTO and academic researchers on the need for the following functions.

- A social model to account for blue-white-red interactions which are of key importance to the success or failure of insurgency warfare.
- A cognitive model to account for red's planning and action decisions.
- Search algorithms/techniques to optimise key parameters for the purpose of identifying good strategies or important interaction phenomena.

This approach leads towards thick agents operating in a simulation environment utilising beliefs, desires and intentions and social network analysis paradigms informed by operational data. Adaptable algorithms such as evolutionary algorithms or other approaches such as simulated annealing will be required to search the complex fitness landscapes that will result from these models/simulations.

There is also consensus that considerable technical, structural/ design, integration, knowledge and expertise challenges exist that must be overcome to achieve the operational CIED CRT capability researchers envision.

²¹ J Millikan, *Computational Red Teaming for Land Experimentation*, presentation to JDSC Computational Red Teaming Workshop, 6/2010

²² Software agents whose behaviour is determined by Beliefs Desires and Intentions that are influenced by the agents actions and interactions with the model environment

²³ N Sayers, *DSTO Operations Support Centre (DOSC) Red Teaming Initiatives 2010*, presentation to JDSC Computational Red Teaming Workshop, 6/2010

4.2 Challenges

The ART inspired approach to CRT shifts the RT process from a creative intuitive exercise towards a modelling and search exercise with the goal of improving the completeness of inquiry of the problem domain.²⁴ Can we do this without unknowingly or unsatisfactorily reducing the dimensions of the problem space – in effect limiting the potential to generate novel insights and solutions? Work exists that provides evidence this is possible, but only for a problem of very modest degrees of freedom²⁵. Modellers and programmers are in effect responsible for providing the scope for creative behaviour in CRT which suggests the need remains for traditional RT approaches leveraging intuitive creative processes within the CRT design phase.

A key challenge will be to manage the issues of verification and validation of complex CRT models. Strategies for managing this will include limiting the scope of modelling and/or documenting the sourcing of data from qualitative methods, and subject matter experts. This highlights the importance of identifying the right problems to be tackled with CRT. It is important to keep in mind that all models are wrong, yet some are useful²⁶. It is also worth noting that RT can generate benefits without formal verification and validation as it is an exercise in problem exploration. Axel Bender's CRT principles²⁷ represent an attempt to provide some boundaries and guidance for the management of these issues.

In order to facilitate any meaningful implementation and impact on military decision making, the client must be committed to the concept of RT, and to structuring their experimentation around this approach. As evidenced below, this requires two key ingredients: the right culture, and high level support.

'The culture of the enterprise: This may be the most important contributor to effective red teaming. Red teaming can thrive in an environment that not only tolerates, but values internal criticism and challenge. Unfortunately, it is often the case that those organisations in need of red teaming have a culture inimical to its use.'

'Top Cover: A red team needs a scope, charter and reporting relationship that fit the management structure. A red team should be expected to raise issues that might not be welcome throughout the enterprise; it needs the support, sometimes from the very top levels of the enterprise. Top cover is needed to ensure that the red team's products not only have the requisite degree of independence, but are seriously considered as well (this does not imply acceptance)'.²⁸

²⁴ H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, Defence and Security Applications Research Centre, University of New South Wales @ the Australian Defence Force Academy, 5/2010

²⁵ C Chua, et al. *Automated Red Teaming: An objective-based data farming approach for red teaming*, GECCO'08, 12/2008

²⁶ Two of the most important and successful theories in physics: Quantum Mechanics and General Relativity are examples of this, as are many superseded theories such as Newton's Laws.

²⁷ A Bender, *Framework for CRT Design*, presentation to JDSC Computational Red Teaming Workshop, 6/2010

²⁸ *The Role and Status of DoD Red Teaming Activities*, Office of the Undersecretary of Defence, 9/2003, <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA430100> accessed 10/2010, pp. 6

UNCLASSIFIED

The services have an established record of RT in their experimentation campaigns. The Army's concept development and experimentation activity known as Headline, incorporates a dedicated multidisciplinary red team. The Navy and Air Force experimentation activities, Headmark and Headway, have also made use of RT. RT could provide substantial benefits to CDG project work and strategic level wargaming such as Force Options Testing and Force Structure Review but these benefits are often not being systematically realised. In order to address this, it is important that the JDSC/JOD works with the client to identify barriers to the uptake and use of RT, and strategies and actions for supporting and encouraging its use. This will be important for preparing the way for future CRT capabilities to play a valuable role supporting client decision making.

It is apparent that with our present levels of knowledge and expertise it would be immature to rush towards challenging applications of CRT such as modelling the CIED domain utilising the CRT approach considered in this report. Knowledge and expertise gaps appear to exist in the following areas that would constrain efforts to produce useful aids to CIED by way of the CRT approach:

- Constructing thick agents and using thick agent tool kits,
- Representation of human factors and social models²⁹, and
- Identifying and applying the best search approaches for the complex fitness landscapes likely to be generated.

It is noteworthy that some of the most serious obstacles encountered have arisen as a consequence of trying to tackle the CIED problem domain which is judged to be inseparable from the fields of social and cognitive modelling. This poses difficulties as social and cognitive sciences are predominantly qualitative sciences that haven't and perhaps cannot be mathematically formulated in an objective and universally accepted manner. This severely limits a researcher's ability to represent social or cognitive phenomena in a rigorous and defensible fashion.

The most sensible approach at current resource levels would be to focus on further research, skill development, and watching progress in key enabling areas for CRT application to CIED. However, with the injection of more resources into this research domain, DSTO/JOD could develop the expertise in JOD staff necessary to exploit software advances as they happen and potentially contribute to the progress of developing the key enablers of CRT towards a more mature state (see section 5). The experimentation community within LOD has certainly voiced support for conducting R&D into the area of sophisticated agents – how to construct them and how to interface them with physical models which flags the potential for meaningful collaboration.

²⁹ A shortage of relevant expertise within JOD/DSTO is a clear limitation though the external academic community may be able to assist.

If more tractable problems were attempted that were better addressed by the skill sets of JOD researchers, it would improve the prospects for demonstrating CRT capabilities in the near term while preparing the way for the integration of additional complexity such as social and cognitive modelling in the future. We can achieve this by developing agent based modelling expertise within JOD which would enhance our ability to represent complex systems numerically. By collaborating with or bringing in social scientists, the challenge of implementing social and cognitive models could be met with lower risk.

The development of agent based modelling expertise would also create the environment and opportunities to learn about the characteristics of the fitness landscapes we must deal with and to select and apply the most appropriate search methodologies to find optimal solutions.

The left field idea of Computation through play/serious games as a CRT approach is worth further investigation. The approach has application to both training and decision support, at the tactical and operational levels particularly with regard to tactics, techniques and procedures. Investigation of the forms this concept could take, and the feasibility and resource demands of potential implementations may yield ways for the ADF to do its business more efficiently and effectively.

4.3 Blind spots

Treatment of adversarial learning which has been highlighted by JOD researchers as an intriguing area worth consideration is largely missing from the academic literature reviews. There appeared to be little consideration of and/or ideas proposed for *in situ* computational tools to enhance RT, or tools to help prepare and facilitate the running of a red team and RT activities. Classified military research and literature in this area have not been reported here. Where do international military research efforts stand in comparison to the picture provided by open source literature?

5. Recommendations

As we move from single service tactical environments to joint operational and strategic scenarios, complexity increases forcing us to abstract the problem. Correspondingly, experimentation moves from verified and validated quantitative numeric models to seminar wargames reliant upon humans bringing the right tacit knowledge to the table on the day and getting the opportunity to raise it. CRT has the potential to improve the quality of investigation of complex problem domains by benefiting military planners and decision makers with useful modelling where previously little to no modelling could be produced.

Below are a number of recommendations from the FY09–10 CRT research program broken up by their specific focus be it CRT as a research field, applying CRT to the CIED domain or CRT in the JDSC.

5.1 General CRT Research and Development

- Demonstrate a proof of principle CRT capability in a simpler problem domain.
- Facilitate JOD CRT researcher involvement in client endorsed RT to better understand the practice, practical issues, utility and limitations of RT in the ADF.
- Support the generation of expertise in and advancement of thick agents, evolutionary algorithms and co-evolution models.
- Investigate how thick agents and co-evolution can be interfaced with existing physical models and identify the challenges to be surmounted.
- Explore the potential of CRT to contribute to Operations Analysis and Operations Research.
- Identify a champion willing to support the implementation of CRT approaches in the client space.
- Investigate the potential utility of serious games and computation through play approaches, possible implementation schemes, their feasibility and likely resource requirements.
- Investigate the relevance of adversarial learning to CRT.

5.2 CIED application of CRT

- Consider the need to bring together social scientists and operational data for the generation of cognitive and social models for application to CIED and counter-insurgency.
- Support the generation of expertise in and advancement of cognitive and social modelling
- Support the CIED threat anticipation modelling efforts underway in the Defence Operations Support Centre; explore its potential role as a test bed or development platform for thick agents, evolutionary algorithms and coevolution.

5.3 Joint Decision Support Centre / JOD specific

- The Joint Decision Support Centre should work with the Capability Development Group to identify barriers to the uptake and use of RT, and strategies and actions for supporting and encouraging its use.

Within the JDSC and JOD, there are a number of disciplines that could benefit from agent based modelling techniques, such as operations research and analysis. A path towards generating the expertise to apply these tools to relevant problems is laid out below:

- Run an in-house course on a standard agent based modelling environment (eg. NetLogo) to raise awareness and expertise (as recommended by the Operations Research Hub in 2009).
- Construct an ABM for analysis of a well-understood/traditional operations research problem.
- Extend the ABM to an initial CRT capability demonstrator.
- Apply CRT to a more difficult problem with the aim of generating insights not easily attainable via alternative means.
- Set up a computer farm for distributed processing of CRT models to reduce computation run times and enable greater complexity to be modelled.

Appendix A: Workshop Attendance

Table A1: Workshop Attendance List

Prof	Hussein Abbass	ADFA	h.abbass@adfa.edu.au
Dr	Michael Barlow	ADFA	M.Barlow@adfa.edu.au
Prof	Peter Campbell	UniSA	peter.campbell@unisa.edu.au
Dr	Mathew Berryman	UniSA	Mathew.berryman@unisa.edu.au
A/Prof	Cara MacNish	UWA	cara@csse.uwa.edu.au
A/Prof	Philip Hingston	ECU	p.hingston@ecu.edu.au
Mr	Phil James	JOD	phillip.james@dsto.defence.gov.au
Dr	Jeremy Manton	JOD	Jeremy.Manton@dsto.defence.gov.au
Dr	Tony Dekker	JOD	Tony.Dekker@dsto.defence.gov.au
Dr	Paul Malcom	JOD/JDSC	Paul.Malcom@dsto.defence.gov.au
Dr	Andrew Gill	JOD	Andrew.Gill@dsto.defence.gov.au
Mr	Nathan Sayers	JOD	Nathan.Sayers@dsto.defence.gov.au
Mr	Justin Millikan	LOD	Justin.Millikan@dsto.defence.gov.au
Dr	Axel Bender	LOD	Axel.Bender@dsto.defence.gov.au
Dr	Daniel Goodburn	C3ID	Daniel.Goodburn@dsto.defence.gov.au
Mr	Duncan Tailby	JOD/JDSC	Duncan.Tailby@dsto.defence.gov.au
Mr	Phillip Gowlett	JOD/JDSC	Phillip.Gowlett@dsto.defence.gov.au

Workshop Attendance	17
---------------------	----

Appendix B: Academic Papers

H Abbass, M Barlow, *Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games*, Defence and Security Applications Research Centre, University of New South Wales @ the Australian Defence Force Academy

P Campbell, *A Literature and Tool Review for IED Computational Red Teaming*, Defence and Systems Institute, University of South Australia, May, **2010**

P Hingston, *Computational Red Teaming – A Literature Survey and Computational Tool Review*, School of Computer and Security Science, Edith Cowan University, April, **2010**

C MacNish, *Computational Red Teaming: A Review for the Defence Science and Technology Organisation*, Faculty of Engineering, Computing and Mathematics, The University of Western Australia, May, **2010**

Computational Red Teaming for Counter Improvised Explosive Devices with a focus on Computer Games

Prof. Hussein Abbass, Dr. Michael Barlow

Defence & Security Applications Research Centre

UNSW@ADFA

Abstract: Red teaming is the art of ethical attack, playing devil's advocate, and learning before acting. The primary purpose of red teaming is risk minimisation. By anticipating behavioural patterns of an opponent, part of the uncertainty is uncovered and the element of a surprise is removed. This chapter describes computational red teaming (CRT), which is an attempt to supplement the red teaming exercise with computational models capable of searching the infinite space of possibilities more efficiently - in cases - than a human would. It then reviews the area of virtual environments and computer games, and their applications to CRT and Counter Improvised Explosive Devices (CIEDs). The chapter concludes with a critical evaluation – through a SWOT analysis - of the suitability of current state of computer games for CRT and recommendations are made on areas of research where investments need to be made to uplift the research area of computer games to be suitable for CRT.

1. Introduction

This report is a compilation of literature, techniques and tools concerning Computational Red Teaming (CRT) as they apply in particular to the challenge of developing doctrine, procedures, tactics, awareness, and technologies for a defence organization facing an opponent who employs Improvised Explosive Devices (IEDs).

The first part of the report provides an understanding of the technique of CRT – its genesis, requisite technologies, current literature, and domains to which it has been applied. The second part of the report concerns an approach to one component of CRT – simulation – that is highly amenable to the IED challenge – that of employing serious military games with human and/or artificial players to explore the search-space of actions and counter-actions for dealing with the IED challenge.

In particular the report has a structure as follows. The next section is a literature review which covers red teaming, CRT, simulation, serious (military) games, and computation through play. That is followed by a section discussing game engines currently employed by defence organizations, as well as open source engines, capable of the appropriate fidelity and resolution for the task. A section is then devoted to an add-on to the VBS2 simulation developed for the US Marine Corp to train their troops in dealing with IEDs (and also understanding the thought processes of those employing them). That is then followed by a section detailing a model of how game play – with the game as the

simulation engine – can be incorporated into the CRT approach. The report concludes with a discussion and set of recommendations.

1.1 History of Computational Red Teaming

Computational Red Teaming (CRT) is a relatively new area in computational modelling. The field of CRT only emerged in 2006 with Australians pioneering the first use of the term and the birth of the field. At that time, a collaboration between the Defence Science and Technology Organisation (DSTO) and the University of New South Wales (UNSW) campus at the Australian Defence Force Academy (UNSW@ADFA) taking the form of a joint supervision of a PhD student, established the first use of Red Teaming (RT) within computational sciences. The thesis commenced with a view to extend the work of the USMC Project Albert by creating an enhanced network centric agent-based distillation capability. The extension resulted in an agent-based distillation system known as the Warfare Intelligent System for Dynamic Optimization of Missions (WISDOM). The thesis then went beyond WISDOM and introduced the seed for computational red teaming by applying fitness landscape analysis to understand combat strategies [Yang, Abbass & Sarker, 2006a]. The Singaporean military then adopted another terminology “automated red teaming” [Choo, Chua & Tay, 2007]. However, their work was simply wrapping an evolutionary computation layer around MANA – another agent-based system from New Zealand that extends the ideas of project Albert. Their line of work was similar to our earlier work [Yang, Abbass & Sarker, 2004, 2005a, 2005b and 2006b]. A formal framework for CRT was introduced in stages in a number of publications including [Abbas 2009, Abbas, Alam & Bender 2009, Abbas, Bender & Whitbread 2010].

It is important to emphasise early in this chapter the difference between CRT and the use of RT in computer science. In the area of network security, RT is becoming a standard technique in testing [Schudel & Wood, 2000], and it falls under what is known as “Penetration Testing”. This use of RT is not different from its traditional use in the military or security domains. However, the ambition of CRT is to undertake RT in Silico. Before we do this, we need to think of RT more as a science and less as an art. Therefore, our formal definitions of RT and CRT are as follows:

Red Teaming (RT) is a structured approach for modelling and executing exercises where reciprocal competitive interaction among two or more players govern the dynamics of how a scenario/situation unfolds with the objective of understanding the space of possibilities, exploring sometimes non-conventional behaviours, testing strategies, and minimising overall risk.

Computational Red Teaming (CRT) is a set of methodologies and computational models that augment a human-based red teaming exercise or perform a computer-based, more abstract, red teaming exercise.

1.2 Fundamental Components of CRT

There are many fields of research that contribute to CRT. While the underlying fundamental and key research questions in CRT may sound to duplicate questions in existing fields of research, the key difference is that questions in CRT stem from the challenges facing an analyst in fusing these fields of research together. It is therefore

imperative to understand the diversity of research areas that contribute to CRT before embarking on activities in CRT that may duplicate research efforts elsewhere. The following figure paints a picture of CRT and the different elements it encompasses.

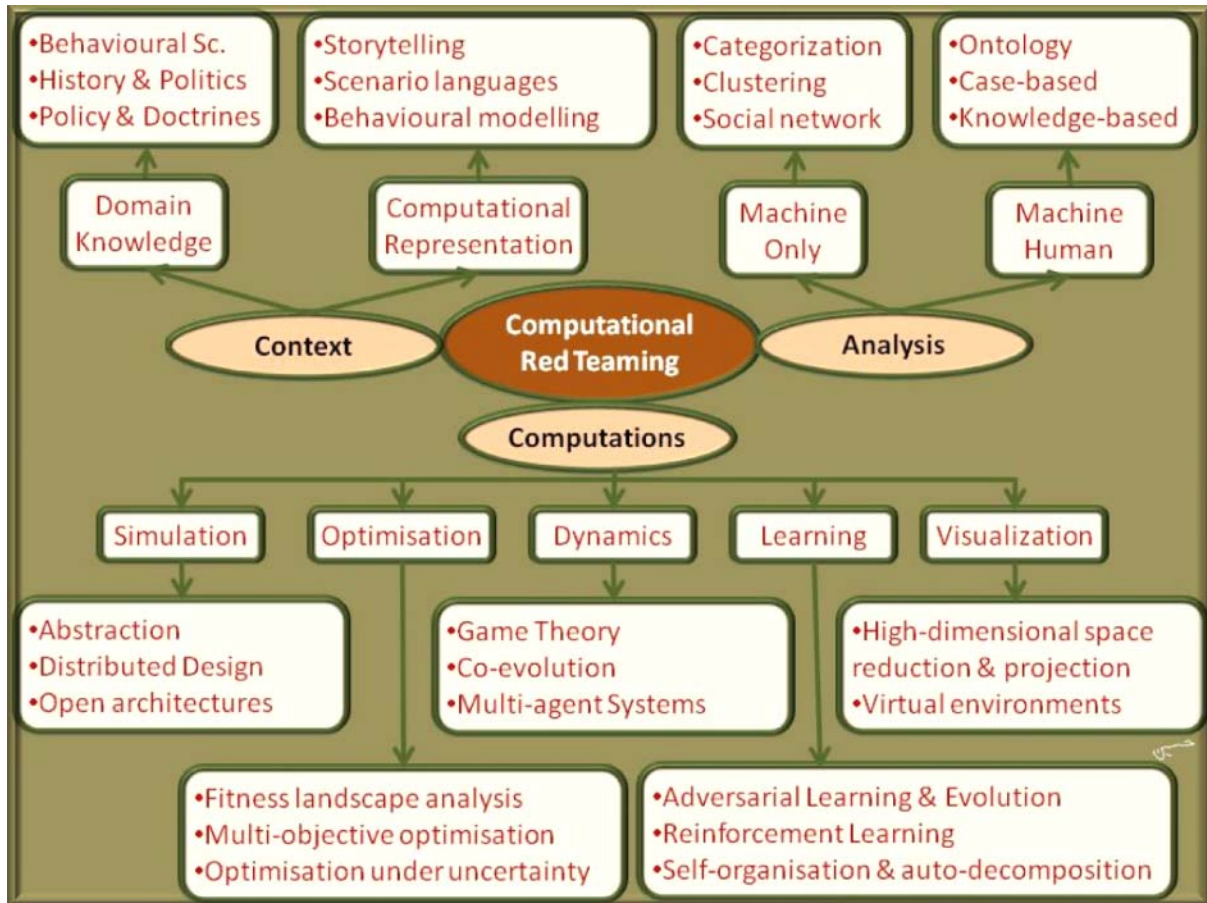


Figure 1: Fundamental components of CRT.

The previous figure depicts the three primary components of CRT; these are: Context, Computations and Analysis. Each of these and their sub-components will be discussed below:

Context: The heart of red teaming is context. It is simply inconceivable to embark on a red teaming exercise without a context. A context provides the meaning behind the activity, the frame of reference, the knowledge required to undertake the exercise, and the environment within which the activity is embodied. The context within CRT can be looked at from two different lenses.

The first is the domain lens, whereby information about the entities to be modelled, their behavioural patterns, political and historical aspects related to the specific scenario to be investigated, policies, procedures, and doctrines that each

of the blue and red teams need to obey are all studied. Domain experts are drawn from all these different fields to contribute to the overall exercise.

The second is the computational lens: how to represent, generate, and manipulate context computationally? The primary issues here include the design of a standard and efficient representation language for a scenario, the use of grammar and constraints to connect events into an overall meaningful story, and the use of efficient representations to capture the behaviour of entities including human. See [Bui, Abbass and Bender, 2010] for an initial study on evolving stories for scenarios.

Computations: Once a context is established, proper representations are designed, the syntax and semantic of the problem are mapped, and scenarios are modelled, the computational side of CRT commences. Here, we need good simulation environments to be able to simulate the dynamics, efficient optimization techniques to mimic human nature in searching for strategies that win, machine learning techniques that can capture the learning elements in blue and red, models of dynamics capable of representing the utilities of different players – be it a rational or irrational process – and visualisation techniques that are able to project the cyber environment onto self-explanatory graphics and motions. Computations in CRT mimics the decision making process of a human team in a real RT exercise. The objective here is not to duplicate human thinking, but at least to not under-represent it in a computational model.

Analysis: At the end of a CRT exercise, one can imagine the myriad of data that are generated from the cyber world. These data need to be grouped, transformed into different formats, and aggregated to higher semantic levels. This is the analysis stage, whereby the myriad of data and information generated throughout a CRT exercise can be represented in a human understandable and comprehensive form.

While CRT leverages all these different areas of research, it would fundamentally be flawed to think of CRT as one or more of these areas, leaving the rest out of the picture. Once more, we need to think of CRT as the glue for all these areas of research to bring them together for the RT exercise.

1.3 Simulation

As identified in Figure 1 above, simulation is a core aspect of the CRT approach – it provides the objective environment in which alternate technologies, tactics, behaviours, doctrines, etc. can be trialed and similarly within which counter-measures (against those technologies and tactics) can be tested and subsequently evaluated.

Simulation is a wide field, and beyond the scope of this report (or even a series of books). Of note though are several issues pertinent to the CRT for IED Challenge task. Firstly, recognizing that simulation is “...an operating model of central features or elements of a real or proposed system, process or environment” (Greenblat, 1988) it is clear that a

(computer) game fits the definition and role of a simulation. Further, the identification of “...central features...” brings in concepts such as abstraction - distilling out non-pertinent elements, fidelity – “trueness” of the model and model components to the portion of the system being modeled, and resolution – to how fine (or low level) a detail is modeled. These are all pertinent for subsequent discussions.

1.4 Games & Serious Games

A multi-billion dollar, international industry, COTS (Commercial Off The Shelf) Computer Games are high-technology mainstream entertainment products. A modern computer game is typically developed by a large team (generally 20 to 50) of individuals – ranging from designers through graphical artists, computer scientists and engineers to musicians and script writers. Most games possess a budget in excess of a million dollars with many now being the order of tens of millions. As a result this highly competitive industry gives rise to hundreds of highly sophisticated and immersive entertainment products each year that often employ the latest in graphical, AI and user interface technology to gain a market edge.

Game genre (see for instance Wolf, 2000), is a broad categorisation of game type on the basis of the game’s subject matter, goal in play, type of interface, player perspective, intended audience and other features. Popular genres include Action – controlling an individual in an interactive 3D environment that typically contains puzzles, fights and other challenges; Sport – football, tennis, etc.; Racing – typically in a car; RPG – roleplaying, typically in a fantasy setting; 1PS/FPS – 1st Person Shooter in which players control heavily armed characters from a 1st person perspective; RTS – Real Time Strategy games; and MMORPG – Massively Multiplayer Online RPG, RPGs with large persistent worlds and a large number of simultaneous players. Certain genres are more commercially successful than others in terms of total revenue or sales; with a difference also existing in fidelity and resolution of the underlying simulation for each genre – hence delineating the suitability of that genre of other non-entertainment purposes.

One offshoot of the commercial “games for entertainment” industry is that known as *serious games* (Prensky, 2000) – the application of entertainment software (“computer games”) or the design and development principles used in that industry, for non-entertainment purposes. Successful examples have included rehabilitation for teen cancer sufferers; UN education about world-hunger, instructions and education on non-violent protest, and training for senior city administrators in dealing with emergencies (Bogost, 2007). The defence domain has been one of the strongest adopters of the serious games approach.

1.5 Serious Military Games

Games technology has considerable utility to support defence organizations in their training, educational, decision-support, and experimental needs. Barlow (2005a), points out that defence organizations ranging from the US Marine Corp, US Army, Australian Army, New Zealand Army, US Navy, US Air Force, through to the UK Ministry of Defence, Danish, Singaporean and Israeli Armies all employ computer game technology for training or experimental purposes.

Amongst the potential benefits can be listed low cost of entry, engagement (Zyda et. al., 2003) and immersion, acceptance (by the new generation of decision-makers), increasing fidelity, multi-player support, ease of modification and “what-if” support, accessibility, and rapid technological advance (Barlow, 2005a). Partially offsetting these benefits are potential risks such as: entertainment genesis of the products, hidden costs, difficulties of on-going support, lack of rigorous testing and validation, and potential for inappropriate “play” (no consequences) attitude by users.

As further evidence of the scale of defence utilization of games, for the last few years NATO has been running a twice-yearly workshop entitled “Exploiting Commercial Games and Technology for Military Use”¹.

1.6 Collective Intelligence & Crowd Sourcing

Collective intelligence is the study of group-level intelligence that emerges as a deliberate or incidental consequence of collaborative or competitive interactions between individual entities within the group. The field has received increasing attention from members of the machine-learning, multi-agent, business, media, and complex systems communities - amongst others. Such attention was initially to note and observe the phenomenon as it existed in real-world systems. However scientific and engineering focus has turned more towards fundamental properties and enablers for the emergence of collective intelligence, together with a methodology for the design of a system (its participants, and their interactions) so that collective intelligence will be guaranteed to arise.

While not a new idea – see for instance Galton’s 1907 paper – Surowiecki is the modern champion of “Crowd Wisdom”, and in his 2004 book, *The Wisdom of Crowds* states: *“Groups do not need to be dominated by exceptionally intelligent people to be smart. Even if most people within a group are not especially well-informed or rational, it can still reach a collectively wise decision. ... [Various statements of the limitations of an individual’s decision-making capabilities] ... Yet despite all these limitations, when our imperfect judgments are aggregated in the right way, our collective intelligence is often excellent.”* Thus, the thesis of Surowiecki’s argument – which he supports with examples drawn from fields as diverse as (horse racing) odds calculation, game shows, and markets – is that an *appropriately constructed* group makes decisions as good as, or better than, experts in that field.

The conditions for a wise decision, and hence by implication a wise-group, as identified by Surowiecki, include the following:

- Diversity – the members of the group should be diverse in that they possess a range of backgrounds, expertise, knowledge-sources, and thought-patterns.
- Independence – group member’s decisions should be independent of one another.
- Decentralisation – Specialisation and usage of local-knowledge should be encouraged. Hierarchical processes and structures should be minimized.

¹ http://idisk.mac.com/dfliesen-Public/nato_presentations/index.html

- Aggregation – A collective decision must be constructed as some form of aggregation of the individual group member's inputs.

Crowd Sourcing, a term originally coined in 2006 (Howe, 2006), but describing an older phenomenon, is an umbrella term for a pool of techniques that at their heart rely on out-sourcing the task of solving a problem to a community - a crowd. Crowd sourcing can be a very haphazard affair - for instance a competition thrown open to the public.

Alternatively, with the recognition of crowd sourcing as inherently a form of collective intelligence, a more disciplined and engineered approach is becoming more the norm.

1.7 Computation Through Play

At the intersection between the serious games movement and crowd-sourcing there is a recent and promising development where game-play serves a computational purpose (von Ahn, 2006).

The Carnegie Mellon academic Luis von Ahn is the most famous exponent of this approach, which is exemplified by his web-based game of “ESP” (Carnegie Mellon University, 2005), which entails pairs of randomly matched players without direct communication means simultaneously labeling images with keywords. Players are “rewarded” (have their score increased) as their keywords match; while, as a result of the play of thousands of people, a labeled database of web images emerges (von Anh & Dabbish, 2004). Other examples of such human-computation games include “Peekaboom” (Carnegie Mellon University, 2005) and the “Google Image Labeler” (Google, 2007). While these games are relatively simplistic they have also been highly successful – both entertaining players (e.g., the google site has a high score board), and resulting in real computation (von Ahn's “ESP” alone has resulted in over 10 million labeled images; there are no figures available for google).

The game designer and futurist Jane McGonigal is a vocal evangelist of designing games and harnessing play to solve real world problems (McGonigal, 2010, 2007a, 2007b). McGonigal's thesis is that current players of computer games are extremely skilled at what they do, and that skill can be employed in the service of solving real-world problems. McGonigal argues that the collective and cumulative power of many millions of hours' play by thousands to millions of individuals can put in service to solve real problems (McGonigal 2007a, 2005) by a deliberate design approach (McGonigal 2007b). In support of this proposition McGonigal has led the design of a number of games such as *Superstruct* which challenges players with an earth of 2019 where a number of potential cataclysms ranging from ecological through terrorism threaten the future of mankind; and *World Without Oil*² which challenged a global community of 1,500 players to deal with the first 32-weeks of a global oil crisis, and transfer those strategies to their real life. Common to these and other games in her stable has been McGonigal's approach of creating an engaged community of players that tackle game versions of very real or potentially real threats.

With regard to game play to answer defence related problems, the Virtual Environments & Simulation Laboratory at UNSW@ADFA has a tradition that extends back to 2001

² <http://www.worldwithoutoil.org/>

(Barlow et. al., 2002a, 2004b). Described in their 2005 paper (Barlow et. al., 2005c) detail a framework where a research question is posed and a game competition is constructed to support the resolution of that question. Scenarios and competition rules are constructed, support infrastructure in terms of equipment but also tools such as match schedulers are provided, participants are enlisted and supported as a community, and data is captured in the form of real-time in-game events and pre/post participation surveys. That data is then analysed and used to inform such questions as how do choices in simulation fidelity influence tactics employed (Barlow & Morrison, 2005b), what factors predict performance of a team (Barlow et. al., 2004a), and how do players make decisions (Lewis & Barlow, 2005).

It is this approach – CRT for the IED challenge achieved as computation through play that is proposed here and elaborated upon in subsequent sections.

2. Open Source & Military Game Engines

Several examples of the successful adoption of games technology and game development approaches for defence needs exist. America's Army³ (Zyda et. al., 2003) is a purpose-built 1st Person Shooter game given away freely by the American Army (through such mechanisms as discs accompanying PC magazines, and web-based download services such as Steam⁴). America's Army is currently in its 3rd edition (version) and is made available as a public relations initiative and recruitment tool. No exact figures are available but it is known that the US Army has spent many millions of dollars in support of America's Army and its ongoing development.

VBS2⁵ (Virtual Battle Space 2) is an Australian developed real-time, 3D simulation product that focuses on the land-domain and for which organizations such as the ADF, UK MoD, US Marine Corp, and US Army possess enterprise licenses. Based on the family of games Operation Flashpoint, Armed Assault, and Armed Assault 2; VBS2 provides a powerful feature set from large detailed external environments, a complete programming language and external scripting interface/API⁶, high fidelity simulation, scenario editor and instructor interface, data capture and replay, through to doctrinal AI, accurate models of modern military units and equipment, and rapid development tools. Considerable academic attention has been focused on VBS (Barlow et. al., 2002a, 2004a, 2005b; Roman & Brown, 2008), including its use in support of defence experimentation (Barlow, Morrison, & Dickie, 2004b). Through its scenario editor and scripting capabilities it is possible to test a hypothesis through designing suitable scenarios, conduct an experiment with suitable participants, and analyse the results of the data capture in-simulation (Barlow et. al., 2004a, 2005b, 2005c). Finally, it is worth noting in this context that VBS2 provides comprehensive in-simulation support for IEDs –

³ <http://www.americasarmy.com/>

⁴ <http://store.steampowered.com/>

⁵ <http://virtualbattlespace.vbs2.com/>

⁶ VBS2Fusion - http://www.simcentric.com.au/index.php?option=com_content&view=article&id=48&Itemid=53

different sizes and trigger mechanisms; ability to emplace, trigger, move etc.; as well as detect and disarm.



Figure 2: Screenshot from the VBS2 3D military simulation.

Finally, Steel Beasts Professional⁷ is a real-time 3D simulation, with its origins as an entertainment product, that focuses on armoured engagement (at the Company level). It is currently or has been used for armoured crew training, as well as education, by armies from Australia, Austria, Canada, Denmark, Finland, Norway, Spain, Sweden, and the USA. Within the ADF, Steel Beasts can and is used in conjunction with VBS2 (and OneSAF) via the mechanism of a HLA/DIS gateway⁸.

Beyond the obvious potential offered by game technology currently utilized by defence organizations, a large untapped pool lies in free game engines. Free game engines – as their name implies – are zero cost game engines. A game engine is (typically) the core simulation and graphical (plus sound, and user interface) software that underpins a game. Sitting on top of a game engine is the content – the 3D models, sound files, dialog, story, etc. Taken together a game engine plus said media make a game. Free game engines vary considerably in their maturity and feature set; with most – but not all – being a generation or more behind the current crop of commercial games. Most free game engines are the work of enthusiastic amateurs, though there are also instances of games companies publicly releasing older (superseded) game engines.

Due to their origins, there is a large, and difficult to track, pool of free game engines available at any time. By way of example Wikipedia, at the time of writing, lists fifty-

⁷ http://www.esimgames.com/steel_beasts_pro.htm

⁸ http://virtualbattlespace.vbs2.com/index.php?option=com_content&task=view&id=92&Itemid=85

one free game engines, and that is a small subset with certain “significant” engines such as FreeCiv⁹ not even listed. Two mature and relevant free game engines are mentioned here. Delat3D¹⁰ is a MOVES institute initiative (originally arising from the America’s Army project, that brings together many open source projects – such as OpenSceneGraph, and Open Dynamics Engine into a single package Darken et. al., 2005; McDowell et. al., 2006). Delta3D has seen application by educational, government, and commercial entities – including defense contractors such as Northrop Grumman. Ogre 3D¹¹ is a graphics (rather than graphics plus simulation) engine that has been used in the development of a large range of graphics, simulation, and game applications. Amongst other usages seen for it, Ogre was employed in a defence experiment to measure the impact of visual representation upon the timeliness and quality of navigation decisions made by maritime officers (Lawes & Barlow, 2007).

3. VBS2 & the Counter-IED Addon

In this section details are presented of an addon for VBS2 that was developed through funds provided by the US Marine Corp. The addon is worthy of attention for several reasons – it concerns IEDs and hence provides an illustration of state of the art simulation fidelity in the representation of IEDs; the addon was developed for training purposes but also addressed the difficult issue of how to deal with the threats of IEDs; and finally the addon took the novel approach of requiring the user to play both sides – be both a convoy leader trying to detect and avoid the threat of IEDs, as well as an insurgent seeking to emplace and inflict the most damage possible with an IED.

This addon, and the development undertaken to produce it, could serve as a guideline or even starting point for the development of a CRT for IED Challenge through play tool.

⁹ http://freeciv.wikia.com/wiki/Main_Page

¹⁰ <http://www.delta3d.org/>

¹¹ <http://www.ogre3d.org/>



Figure 3: A screenshot of the Counter-IED VBS2 add-on. The shot illustrates the Marine convoy leader's interface, including the command interface in the lower-right (note ability to order AI to conduct "5 and 25" or mount up, etc.). A suspected IED has been detected by one of the AI and its location is indicated by the arrow.

In 2008 Cognitive Training Solutions (CTS), a US firm, received funding from the US Marines to build a "Counter-IED" trainer. CTS partnered with Bohemia Interactive Australia (BIA) – developers of VBS2 – to develop an integrated solution, the main component of which was an add-on for VBS2 (Bohemia Interactive, 2009b). BIA put together a team of developers to bring the add-on to fruition, with this task being project managed by one of this report's authors. The add-on was delivered in early 2009 and subsequently included in some later versions of VBS2 – such as *VBS NATO* (Bohemia Interactive, 2009a).

CTS sought to deliver a solution with a cognitive focus, with a core concept being 'Insurgent Mindset Training' – that is the marine participant would experience both sides and become cognizant of the features and cues that an insurgent would employ in obtaining their goals.

To support such a goal, considerable development was required ranging from additional 3D models, through core engine changes, alterations to AI behaviours (a significant component) to new interfaces, media, and scripting changes.

A suite of some dozen-odd single-player scenarios was developed. In half the scenarios the player led a convoy of HMMVs (and at times other vehicles), navigating across the map (urban or desert) and seeking to arrive safely at the designated destination by avoiding and/or detecting (and having disposed of) any IEDs along the path. In these cases IED placement was machine controlled and pseudo-random (chosen from a large pool of pre-identified ambush sites and configurations) with any ambushes upon the convoy carried out by scripted AI. Similarly the player led a group of AI avatars – the drivers and occupants of the vehicles – that could be ordered to debus, conduct sweeps and searches, etc.

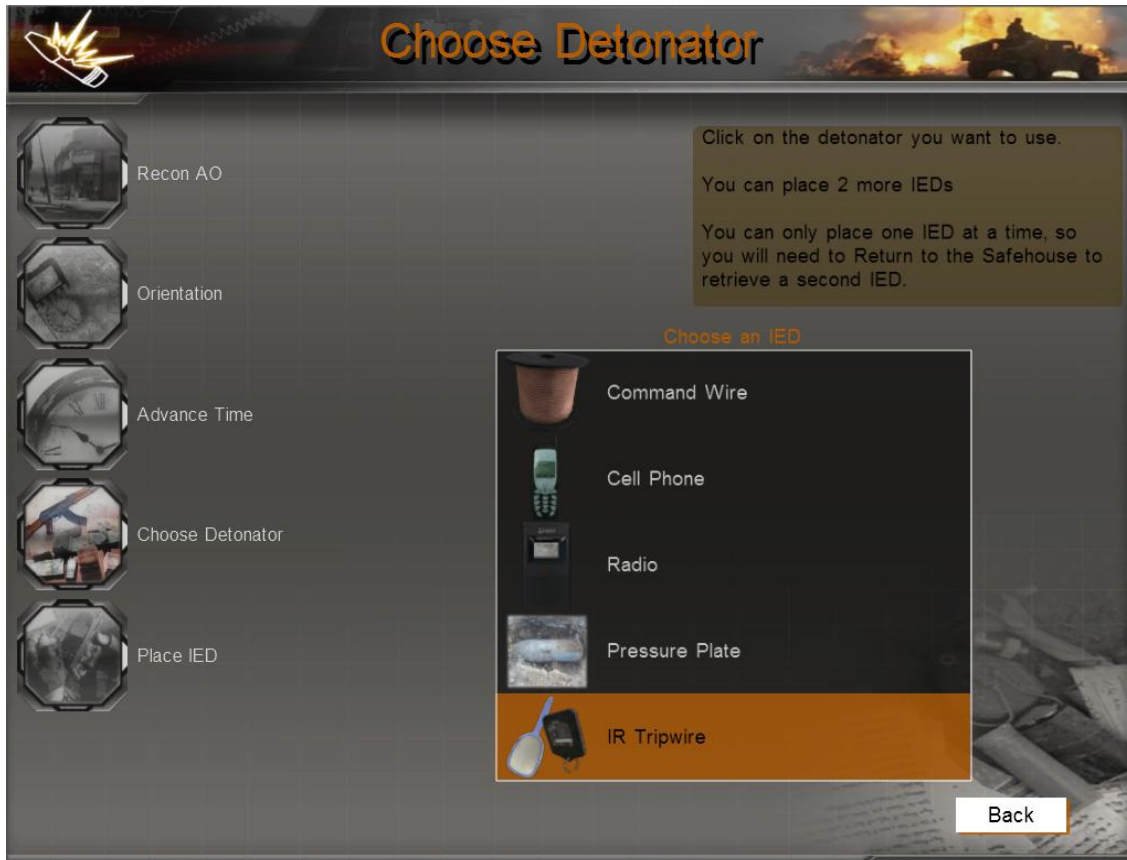


Figure 4: A screenshot of the Counter-IED VBS2 addon. The shot illustrates the insurgent's menu-system, available to them when they occupy the safe-house. Through this menu system they could choose IED and trigger devices, receive a briefing, advance time, and choose to leave to reconnoitre an area or place the IED.

In the other half of the scenarios the participant played as an insurgent with differing assets at their disposal to conduct an IED ambush on a marine convoy. In these scenarios the insurgent operated out of a safe-house, starting twenty-four hours or more before the convoy was due along its route, and through a menu system (see Figure 4 for an example) was able to control the advance of time so as to choose appropriate times of day to scout an area, or emplace an IED. Again scenarios occurred in urban and desert roadside settings, with the player able to choose the exact location of the IED, their own location for the ambush, etc. In these cases the marine convoy was AI controlled and acted

according to scripted doctrine, while any insurgent accompanying the player (provided in some scenarios) were under AI control but could be ordered about by the player.



Figure 5: A screenshot of the Counter-IED VBS2 add-on. The shot illustrates the insurgent's viewpoint as an ambush is unfolding – the insurgent is upon the roof of a building firing upon the members of the convoy that has stopped in response to an IED detonation.

Features of VBS2 and the Counter-IED add-on, pertinent for any potential CRT-IED gaming approach include:

- Ability to play “both sides” of the conflict either as single-player, cooperative (multi-player, all on same side) or full multi-player (multiple players on each side)
- Significant range of IED representation and in-simulation support for trigger mechanisms ranging from command-wire to pressure plate, IR beam, and cell-phone.
- AI support for convoy members, insurgents, and civilians (considerable work was done on civilian behaviours so they reacted appropriately to the dynamic environment and provided cues that an observant convoy leader could detect).

- Simulation support for electronic counter measures to IED triggers (such as jammers).
- Data-capture of all scenario runs that was extended to an automated scoring and feedback system.
- A C++ API allowing for external AI, simulation or other computations in a general purpose language.

It appears apparent that the addon described above “captures” a lot of the tactical considerations surrounding IEDs, their usage, and counter-measures. Coupled with VBS2’s feature set of scenario-editors, large-scale urban and rural environments, AI, data capture, and instructor interfaces, it would appear a highly suitable starting point for a CRT-IED task that focused at that tactical level on issues such as doctrine, tactics, and equipment.

4. Games for CRT - Details & Issues

When using games for CRT, certain broad principles apply and have been discussed elsewhere (Barlow, 2005a). Of particular interest, the game engine already provide a relatively high (suitable) fidelity of the environment (e.g., tactics and counter-tactics to the usage of IEDs) to be explored; that the game be modifiable (so as to meet the actual requirements); and that some form of data-capture facility exist to enable analysis and review.

When a game and play approach is suggested this should be understood in the broadest sense. That is the game is not simply the particular software but the entire environment that engages players. As a crowd-sourcing approach is suggested this means that the activity from the player’s perspective must be interesting and engaging. This means that not only must be play of the game itself be enjoyable, but that the surrounding activities that support the activity (e.g., player enrolment procedures, pre and post-game surveys) be constructed so as to maintain participant engagement and enthusiasm. This can often entail the construction of supplementary resources – scoreboards and feedback for players to engage their competitive spirit, web sites and discussion boards to support the community and its sense of self and contribution. There is a considerable literature within the games design area (e.g., Swink, 2009; Schell, 2008) concerning this holistic approach to game design.

With a crowd-sourcing approach, the question of the crowd composition – its size and the skills and background of its individual members is something that can be addressed with an appropriate design. That is, a networked game offers the potential to engage thousands of players; but how many are necessary for a particular problem, and should they be selected with a particular skillset in mind – perhaps serving ADF members for an IED task, or should Surowiecki’s (Surowiecki, 2004) assertion of the wisdom of a truly random or “unskilled” crowd be trusted?

With the prospect of a standard game’s support for human only, human with AI, and even AI only (not supported by all), different options with regard to exploration of the decision-space are offered. While human alone offers the prospect of exploration from a human decision-making perspective – at the higher resource cost of the participants’

time; AI only allows larger scale as well as greater breadth of exploration to be performed for the same resource cost. A hybrid approach of one or more human players controlling or interacting with multiple AI controlled entities appears to offer a compromise where the best features of both approaches might be extracted.

Perhaps more fascinating is the concept where AI and human participants are mixed through time rather than in a single scenario. For instance, an evolutionary computation technique might be employed to run through thousands of scenarios with AI only participants. When an interesting new tactic appears to have emerged this can be tested and validated with human participants exploring the same approach and its counter tactics. The results of the human trials might then be fed back into the next round of the evolutionary process. Such an approach of swapping between humans and AI pursuing the same in-simulation stratagems is only possible with something like a game and its support for human or AI control of avatars. Clearly computational resource challenges exist for this concept of AI exploration of decision-space. On the other-hand, in a distributed/networked game, computational resources expand as players join and their game platform becomes available. Approaches such as SETI@home¹² (Sullivan et. al., 1997) and folding@home¹³ (Shirt & Pande, 2000) - approaches which use the idle “screen saver” cycles of people’s PCs to solve protein folding and signal processing problems – provide models of how this might be achieved.

Fascinating questions also exist concerning guidance for and over-sight of the process once initial results start to emerge. A powerful but challenging approach is to view the process as dynamic – results from early scenarios and runs informing subsequent scenarios, with those subsequent scenarios and runs being adjusted or created in response to the emerging early results. There are a large number of design challenges here – how to design the entire process to support such a dynamic and agile model; how to achieve the analysis and turn-around in new scenario creation in a timely fashion; how to ensure on-going engagement with the player-pool as the game scenarios go through these iterations; and even technically how to propagate out such changes to all players.

Finally, in that vein there are a number of practical issues and challenges. How is the game distributed or made available to players; how is the player-pool managed including registration/enrolment, and verification/validation; how is data capture handled in such a distributed environment; and what resources are required to create, then conduct and maintain the tool, as well as to perform the analysis on the resulting data? Other, perhaps deeper, but still practical considerations also exist: If simulation fidelity needed to be altered, how could this be supported (a scripting language – see Barlow & Morrison, 2005b for an example – or even better an API might facilitate or at least ameliorate this)?

5. CRT for IEDs

While the focus in a number of CIEDs activities has been placed on prediction, the potential for applying CRT to CIEDs is more general than prediction alone. CRT can be used in two forms:

¹² <http://setiathome.berkeley.edu/>

¹³ <http://folding.stanford.edu/>

As an educational and training tool: most of the work in the literature reduces to a simple simulation environment, whereby officers get exposed to a number of pre-scripted scenarios allowing some interactivity from the user end, to deposit IEDS, predict, detect and possibly prevent enemy activities. Alternatively, an environment is made available to allow both red and blue to red team. The VBS2 Counter-IED add-on described above falls in both categories. CRT should offer this literature much more than a game-type environment. Take for example intervention methods; this can range from a psychological influence of a political will to high-tech robotic technologies. A CRT environment for CIEDs education and training would require the level of computational sophistication in Figure 1 to truly interact and respond effectively to a user, go beyond scripted AI to modelling the ability to evolve and adapt, understanding the driving forces of the dynamics of the game, analysing the interaction, and synthesizing a large number of scenarios into proper response surfaces.

As a planning what-if style system. Work in this area has only been on the tactical level. Going above a tactical level – and especially if we need to link tactical to operational and strategic levels – would require the level of sophistication discussed above to be able to detect change, measure and aggregate performance indicators, decompose and aggregate strategies and plans, to name a few.

In general, one way to imagine the different functions that a CRT framework would undertake for CIEDs is the following:

Anticipation: this function can be translated computationally into a set of heterogeneous data sources that need to be fused, processed, and semantically correlated with a frame of reference to inform the decision making process with an appropriate level of situation awareness and a prediction with a reasonable level of confidence.

Detection: this function is concerned with the set of processes, procedures and doctrines that need to be in place to detect IEDs in situ. The detection problem is tightly coupled with the planning for troop movements, formations, and logistic support. It also incorporates search patterns, the use of different technologies such as robots, and application of different counter measures.

Prevention: this function is concerned with the interception with the red-team planning process, denying their capabilities, disrupting their supply chain, reshaping the environment, and deceiving them. Avoidance is another prevention strategy whereby places with a high level of threat, for example, can be avoided.

Protection: this function is normally done non-computationally. It produces technologies that can protect blue such as protected vehicles and special materials worn to CIEDs.

Response, recovery and neutralization: this function is concerned with post the attack stage. Strategies, procedures and doctrines are designed and red teamed to mitigate damage and expedite the cleanup process.

The previous discussion outlines the importance of CRT and its potential use for CIEDs.

6. A SWOT Analysis of the Suitability of Games for CRT

In the previous sections, we discussed serious games, military games, collective intelligence and computation through play. In this penultimate section, we will group these areas under one title: Games. A Strengths, Weaknesses, Opportunities, and Threats analysis (SWOT) will be conducted to evaluate the use of games for CRT. The basis of our argument is that a game can be employed to provide the simulation.

6.1 Strengths

- S1. [**Suitable for IED modelling**] Certain games appear to have the richest and highest fidelity representation of IED effects, the environments in which IEDs are employed, and the tactics and counter-tactics surrounding their usage. Even if a low-tactical level simulation of the IED environment (such as offered by VBS2 or similar tools) is not desired, another genre of game (e.g., real-time-strategy or turn-based) is likely to be able to capture the emphasis and decision-space required.
- S2. [**Harnessing creativity**] Engaging humans in the CRT framework through play of the game harnesses the creativity and inventiveness of human thought to explore the decision-space in an efficient and novel manner, while also face-validating any approaches.
- S3. [**Capturing the wisdom of the crowd**] A networked game allows a true crowd-sourcing approach – where hundreds, thousands, tens of thousands or more players can robustly test and explore ideas. Due to their engaging and immersive nature it is possible to “capture” this many participants simply through the entertainment and enjoyment they obtain via play.
- S4. [**Mixing AI and Human**] Game support for AI, human, or mixed AI and human players, and relatively realistic behaviours by AI, allow for tuning of the CRT tool along dimensions of higher fidelity, larger scale, and rapid exploration.

6.2 Weaknesses

- W1. [**A fixed fidelity model**] Choice of the core game engine is obviously a task dependent issue (e.g., the most suitable game engine for tactical considerations concerning IEDs may not be the best for one regarding logistics, or regarding medical issues arising from IED attacks, or ...) but this entails that each time a task changes, either the game needs to be re-developed or a new game needs to be adopted. In almost all computer games, the game designer selects a fixed level of fidelity of the world being modelled. One can't change this level of fidelity within a game. It is not even easy to reprogram a game with a different level of fidelity. When a game idea emerges, there is a long lead time between the idea and the creation of the game.
- W2. [**High computational cost**] In almost every computer game available in the market, there is an inherent characteristic that the complexity of the game exceeds with orders of magnitude the complexity of the actual task. Imagine the Prisoner Dilemma (PD) theoretic game. To create a computer game for PD, it needs to be interesting with nice visualization, engaging and entertaining. However, all these features that we add to make the game more interesting for a human deviate from the core underlying dynamics; that is, the PD game. Moreover, they introduce

elements of variations that are not important in analyzing the dynamics of the PD game.

- W3. [**Absence of explicit models and Immature AI**] Unfortunately, the game industry does not separate the model from the implementation. The mathematical basis for the model underlying a game is entangled with the code and is maximally inseparable. The AI implemented in commercial games nowadays is scripted and sometimes one would wonder if these scripted agents should be called AI at all. Attempts have been made to use reinforcement learning, neural networks and classifier systems to control the AI within some game environments such as Second Life. However, this line of research is still in its infancy.

6.3 Opportunities

- O1. [**A multi-billion dollar industry**] There are opportunities to tap into the game industry and influence the evolution of this industry for CRT in a manner similar to how military training influenced it.
- O2. [**An all-age exercise**] Games are played by people from different age-groups, not just children, and this extensive & diverse range of skills and backgrounds can be tapped-into for CRT-purposes via games.
- O3. [**Multiple Tools**] The games industry produces hundreds of games each year in a range of genres (from highly immersive 1PS through to abstract puzzlers) across a range of hardware platforms. At the same time there is a burgeoning “indie” development industry as the tools for development become simpler to employ (and game development skills continue to spread).

6.4 Threats

- T1. [**A hard to control market**] The computer game industry is particularly large, with an inertia of its own and a number of significant players (e.g., companies like Microsoft or Sony) already seeking to exert control and influence over the industry’s development and future.
- T2. [**Replacing simulation with games**] Games have their origins in the desire to entertain rather than a desire to create an objective simulation of (some) reality. The consequences being differences – subtle and otherwise – can exist between a game/entertainment simulation and an objective simulation.
- T3. [**Rapid change in technology**] The games industry is highly competitive, with developers and producers always pursuing a competitive-edge that will allow them to differentiate and promote their product. One consequence is that new technologies are rapidly being adopted...and often as rapidly abandoned.

UNCLASSIFIED

The following table lists the possible actions that need to be taken to address the points made in SWOT.

		Opportunities	Threats
		O1. A multi-billion dollar industry O2. An all-age exercise O3. Multiple Tools	T1. A hard to control market T2. Replacing simulation with games T3. Rapid change in technology
Strength	S1. Suitable for IED modelling S2. Harnessing creativity S3. Capturing the wisdom of the crowd S4. Mixing AI and Human	A1: O1,S1 quickly move in this direction A2: O2,S2,S3 develop a distributed web-based game A3: O2, O3,S4 allow multiple levels of complexity	A4: T1,S1,S3 create the demand for CIEDs games A5: T2,S2,S4 design the game on simulation principles A6: T3,S1 separate the model from the game and both from the visualization
Weakness	W1. A fixed fidelity model W2. High computational cost W3. Absence of explicit models and Immature AI	A7: O1,W2 use grid computing in the design A8: O2,W1 allow multiple levels of complexity A9: O3,W3 articulate clear requirements and specifications to include AI	A10: T1,W1 allow flexible multi-level of resolutions A11: T2,W3 separate the model from the game and both from the visualization A12: T3,W2 use service oriented architecture design and grid computing

The above analysis highlighted 12 different actions; these are:

- A1: quickly move in this direction
- A2: develop a distributed web-based game
- A3: allow multiple levels of complexity
- A4: create the demand for CIEDs games
- A5: design the game on simulation principles
- A6: separate the model from the game and both from the visualization
- A7: use grid computing in the design
- A8: allow multiple levels of complexity
- A9: articulate clear requirements and specifications to include AI
- A10: allow flexible multi-level of resolutions
- A11: separate the model from the game and both from the visualization
- A12: use service oriented architecture design and grid computing

UNCLASSIFIED

Actions A1 and A4 are related to the opportunity and could be combined into: create the demand and move quickly to adopt games for CRT.

Actions A2, A3, A7, A8, A10, A12 are related to the game design and could be combined into the following design specifications: service oriented architecture, multi-level of resolutions, web-based, distributed design, and grid computing design.

Actions A5, A6, and A11 are related to modelling and could be combined into: ensure a complete separation between the model underlying the game and the actual implementation of the game including the GUI.

Action 9 stands on its own right as the need for much more advanced AI.

6. Discussion & Recommendations

In this chapter, we laid out the history of CRT, reviewed computer games and their use for CIEDs. The chapter concluded with a critical assessment of the state-of-art of computer games and its suitability for CRT.

The Strength, Weakness, Opportunity and Threat (SWOT) approach revealed the challenges that we need to overcome to apply computer games to CRT. The SWOT analysis ended up with four clear recommendations:

Recommendation 1: There is an urge for creating a demand for the use of games for CRT. This move needs to be taken quickly before the market becomes uncontrollable.

Recommendation 2: The game to be designed for CRT should possess a number of design principles and features including service oriented architecture, multi-level of resolutions, web-based, and distributed design. It should also utilise the grid computing technology.

Recommendation 3: In the design of games for CRT, there should be a clear separation between the model underlying the game and the actual implementation of the game including the GUI. This separation make the game more suitable as a decision making model and not merely as a source of entertainment. It also reduces the risk resultant from technology change.

Recommendation 4: Games for CRT should have much better AI to ensure their suitability for searching the space of possibility in CRT.

References

Abbass H.A. (2009) Computational Red Teaming and Cyber Challenges, Invited Speech, Platform Technologies Research Institute Annual Symposium, PTRI 09.

Abbass H.A., Alam S., and Bender A. (2009) MEBRA: Multiobjective Evolutionary Based Risk Assessment, IEEE Computational Intelligence Magazine, IEEE Press, Pages 29-36.

Abbass H.A. and Bender A. (2009) The Pareto Operating Curve for Risk Minimization, Artificial Life and Robotics Journal, Vol.14.

- Abbass H.A., Bender A., and Whitbread P. (2010) Computational Red Teaming: Unravelling the Pharaoh's Curse, Computational Intelligence in Security and Defence Applications Workshop, WCCI 2010, Barcelona, Spain.
- Alam S., Shafi K., H.A. Abbass and M. Barlow (2009) An Ensemble Approach for Conflict Detection in Free Flight by Data Mining, *Transportation Research Part C*, 17(3), Pages 298-317
- von Ahn L., Dabbish L. (2004) "Labeling Images with a Computer Game", *Proc. Of the SIGCHI Conf. on Human Factors in Computing Systems*, 319-326.
- von Ahn L. (2006) "Games with a Purpose", *IEEE Computer*, Vol 39, 6, 92-94.
- Barlow M. (2005a) "The Game of Defence & Security", in *Applications of Informations Systems to Homeland Security and Defence*, Abbass & Essams (Eds), pages 138-166, Idea Group Inc., USA.
- Barlow M., Morrison P. (2005b) "Challenging the Super Soldier Syndrome in 1st Person Simulations", *proceedings of SimTecT 2005*, pages 13-18, Sydney Australia, May 2005.
- Barlow M., Luck M., Mihe A., Morrison P., Lewis E. (2005c) "A Novel and Sustainable Infrastructure for Virtual Environment Research", *proceedings of SimTecT 2005*, pages 249-254, Sydney Australia, May 2005.
- Barlow M., Luck M., Lewis E., Ford M., & Cox R. (2004a) "Factors in Team Performance in a Virtual Squad Environment", *Proceedings of SimTecT 2004*, 94-99, Canberra, May 2004.
- Barlow M., Morrison P. & Dickie A. (2004b) "Constructing the Virtual Section", *Proceedings of SimTecT 2004*, 100-105, Canberra, May 2004.
- Barlow M., Morrison P. & Easton A. (2002a) "1st Person Tactical Shooters: COTS Games with Military Training Potential", *Proceedings of SimTect2002*, Melbourne, May 2002.
- Barlow M., Ford M., Lewis E. (2002b) "The Use of Commercial Off the Shelf Games by Military Forces", *Report for Australian Defence Simulations Office*, June.
- Barlow M., Ford M., Lewis E. (2002c) "The Use of Commercial Off the Shelf Games by Military Forces [Review of Games]", *Report for Australian Defence Simulations Office*, June.
- Bogost, I (2007) *Persuasive Games: the Expressive Power of Videogames*, MIT Press
- Bohemia Interactive (2009a) "NATO to deploy VBS2 for tactical training", *Press Release*, http://virtualbattlespace.vbs2.com/index.php?option=com_content&task=view&id=119&Itemid=73, accessed 8/4/2010.
- Bohemia Interactive (2009b) "VBS Newsletter – January 2009", http://vbs2.com/media/newsletters/09-01_VBS_Newsletter.htm, accessed 8/4/2010.
- Bui L., Abbass H.A., Barlow M., and Bender A. (2010) Robustness Against the Decision-Maker's Attitude to Risk in Problems with Conflicting Objectives, *IEEE Transactions on Evolutionary Computation*, accepted 6 April, 2010, In Press

Bui L.T., Barlow M. and Abbass H.A. (2009) A Multiobjective Risk-based Framework for Mission Capability Planning, New Mathematics and Natural Computation, World Scientific, 5(2), pp 459-485.

Bui V., Abbass H.A. and Bender A. (2010) Evolving Stories: Grammar Evolution for Automatic Plot Generation, IEEE Congress on Evolutionary Computation, Barcelona, Spain.

Carnegie Mellon University (2005) "The ESP Game: Labeling the Web", <http://www.espgame.org/>, accessed 7/1/2008.

Carnegie Mellon University (2005) "Peekaboom: Not just wasting your time", <http://www.peekaboom.org/>, accessed 7/1/2008.

Choo CS, Chua CL, and Tay SHV (2007) Automated Red Teaming: A Proposed Framework for Military Application, GECCO 2007.

Dam H., H.A. Abbass, Lokan C. and Yao X. (2008) Neural-Based Learning Classifier Systems, IEEE Transactions on Knowledge and Data Engineering, vol 20(1), 26-39.

Darken R., McDowell P., Johnson E. (2005) "Projects in VR: the Delta3D open source game engine", IEEE Trans. On Computer Graphics and Applications, Vol 25, #3, pages 10-12.

Galton F. (1907) "Vox populi", Nature, Vol 75, pages 450-451.

Ghoneim A., Abbass H.A. and M. Barlow (2008), Characterizing Game Dynamics in 2-Player Strategy Games Using Network Motifs, IEEE Transactions on Systems, Man, Cybernetics, Part B, vol 38(3), 682-690

Google (2007) "Google Image Labeler", <http://images.google.com/imagelabeler/>, Accessed 7/1/2008.

Greenblat C. (1988) *Designing Games and Simulations – An Illustrated Handbook*, London: Sage Publications.

Howe J. (2006) "The Rise of Crowdsourcing", Wired, 14, June 2006, <http://www.wired.com/wired/archive/14.06/crowds.html>, accessed 7/4/2010.

Lawes G., Barlow M. (2007) "Visual Realism and Decision Making: A Novel Approach to Real-Time Maritime Battlespace Visualisation", SimTect07

Lewis E., Barlow M. (2005) "The Use of Games to Investigate Tactical Decision-making, proceedings of SimTecT 2005, pages 73-78, Sydney Australia, May 2005.

Michael, D., and Chen S. (2005) *Serious Games: Games that Educate, Train, and Inform*, Course Technology PTR.

McDowell P., Darken R., Sullivan J., Johnson E. (2006) "Delta3D: A Complete Open Source Game and Simulation Engine for Building Military Training Systems", Journal of Defense Modeling & Simulation, Vol 3, #3, Pages 143-154.

McGonigal J. (2010) "TED Talks – Jane McGonigal: Gaming can Make a better world", <http://www.youtube.com/watch?v=dE1DuBesGYM>, accessed 7/4/2010.

McGonigal J. (2007a) "Why I Love Bees: A Case Study in Collective Intelligence Gaming", in *The Ecology of Games: Connecting Youth, Games and Learning*, pages 199-227, MIT Press.

McGonigal J. (2007b) "The Puppet Master Problem: Design for Real-World, Mission-Based Gaming", in *Second Person: role-playing and story in games and playable media*, eds. Pat Harrigan & Noah Wardrip-Fruin, MIT Press.

McGonigal J. (2005) "Supergaming! Ubiquitous Play and Performance for Massively Scaled Communities", *Modern Drama*, Vol 48, #3, pages 471-491.

Morrison M., Barlow M., Bethel G., Clothier S. (2005) "Proficient Soldier to Skilled Gamer: Training for COTS Success", proceedings of SimTecT 2005, pages 91-96, Sydney Australia, May 2005.

Morrison P. & Barlow M. (2004) "Child's Play? Coercing a COTS Game into a Military Experimentation Tool", Proceedings of SimTecT 2004, 72-77, Canberra, May 2004.

Nguyen M.H., Abbass H.A. and McKay R. (2008) Analysis of CCME: Coevolutionary Dynamics, Automatic Problem Decomposition and Regularization, IEEE Transactions on Systems, Man, Cybernetics, Part C, vol 38(1), 100-109.

Prensky M. (2000) "Digital Game-Based Learning", McGraw Hill.

Quek H.Y., Tan K.C., and Abbass H.A. (2009) Evolutionary Game Theoretic Approach for Modeling Civil Violence, IEEE Transactions on Evolutionary Computation, 13(4), Pages 1-21.

Quek H.Y., Tan K.C., Goh C.K., and Abbass H.A. (2009) Evolution and incremental learning in the iterated prisoner's dilemma, IEEE Transactions on Evolutionary Computation, 13(2), 303-320.

Rojanavas P., Dam H.H., Abbass H.A., Lokan C. and Pinngern O. (2009). A Self-Organized, Distributed, and Adaptive Rule-Based Induction System, IEEE Transactions on Neural Networks, 20(3), 446-459.

Roman P.A., Brown D. (2008) "Games – Just How Serious Are They?", Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2008

Schell J. (2008) *The Art of Game Design: A book of lenses*, Morgan Kaufmann.

Schudel and Wood, DARPA IA Red Team Experiments, MILCOM 2000, IEEE Publisher.

Shirts M., Pande V.S. (2000) "Screen Savers of the World Unite!", *Science*, Volume 290, Number 5498, Issue of 8 Dec 2000, pp. 1903-1904.

Sullivan W.T., Werthimer D., Bowyer S., Cobb J., Gedye D., Anderson D. (1997) "A new major SETI project based on Project Serendip data and 100,000 personal computers", in "Astronomical and Biochemical Origins and the Search for Life in the Universe", Proc. of the Fifth Intl. Conf. on Bioastronomy.

Surowiecki, J. (2004). "The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations", Little, Brown

- Swink S. (2009) *Game Feel A Game Designer's Guide to Virtual Sensation*, Morgan Kaufmann Publishers, Burlington, MA, USA.
- VESL (2003) "Report on the Headline03 Virtual Infantry Section Experiment", VESL Report 2003-1, November.
- Wang S.L., Shafi K., Lokan C. and Abbass H.A. (2010) Robustness of Neural Ensembles Against Targeted and Random Adversarial Learning, IEEE International Conference on Fuzzy Systems, Barcelona, Spain.
- Wolf M. J. P. (2000) "Genre and the video game", in *The medium of the video game*, University of Texas Press.
- Yang A., Abbass H.A., and Sarker H (2006a). Characterizing Warfare in Red Teaming, IEEE Transactions on Systems, Man, Cybernetics, Part B, 36(1)
- A. Yang, H.A. Abbass, and R. Sarker (2006b) Land Combat Scenario Planning: A Multiobjective Approach, 6th International Conference on Simulated Evolution and Learning (SEAL'06), LNCS 4247, Hefei, Chinapp. pp. 837-844.
- Yang A., Abbass H.A., and Sarker R. (2005a). WISDOM-II: A Network Centric Model for Warfare. Ninth International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES 2005), Lecture Note Computer Science LNCS 3683, Melbourne, Australia, pp. 813-819.
- Yang A., Abbass H.A., and Sarker R. (2005b) Risk Assessment of Capability Requirements Using WISDOM-II, The International Society for Optical Engineering (SPIE), Microelectronics, MEMS, and Nanotechnology Symposium, Complex Systems Conference, Brisbane, Qld.
- Yang A., Abbass H.A., and Sarker R. (2004) Landscape Dynamics in Multi-agent Simulation Combat Systems", The 17th Australian Joint Conference on Artificial Intelligence, Lecture Notes in Artificial Intelligence, LNAI 3339, Springer-Verlag, 39-50.
- Zyda M., Hiles J., Mayberry A., Wardynski C., Capps M., Osborn B., et. al. (2003) "The MOVES Institute's army game project: Entertainment R&D for defence", IEEE Computer Graphics and Applications, Jan/Feb



A Literature and Tool Review

For IED

Computational Red Teaming

1. Introduction

The idea of understanding the future actions and reactions of one's adversaries goes back many centuries (Sun Tzu ~500 BCE). In more modern times war gaming has been employed in many different ways with many of the large scale games including blue, green, white and other colours of teams representing various players, as well as red teams representing the main adversary. A practical definition of the Red Teaming process for conventional warfare might be as follows: *An iterative, interactive process conducted during CAP to assess planning decisions, assumptions, COAs, processes, and products from the perspective of friendly, enemy, and outside organizations* (Malone 2002). The US Defense Science Board provides the following definition-“*Red teaming* is defined as an activity using a person (or group of people)—sometimes as adaptive simulated enemies—to look for and test vulnerabilities in military plans and/or emerging technical concepts.” (Defense Science Board, 2003.)

In the case of counter IED red teaming there are two major application areas that must be considered. Firstly there is the application corresponding to the use of red teaming in conventional warfare – that is, capability assessment planning (CAP) where the goal is to understand and plan adversary reaction to blue force initiatives, responses to longer term trends and resource constraints, technology changes and the many other issues that will change over the course of an

operation that might last for years. The second application is operational red teaming to support day to day operations where the operational tempo is very high. Here the goal is to use whatever information is available from field reports, contemporary media reports, intelligence information about current technology use by terrorists, information gained from the local populace, local and theatre geo-spatial information and any new blue defeat technology that may become available to help day to day mission planning. This help may take several forms, including selection of appropriate counter measures, probable location of IEDs, subtle indicators of possible IED related activity around blue force patrol routes, mid-term (one to two weeks) prediction of possible IED attack patterns and randomising of route selection for blue force activities within the context of the situation assessment.

In this paper, Computational Red Teaming is taken to mean the use of computer models and/or other mathematical techniques, using a variety of modelling approaches, that provide support to the red team in exploring and understanding the complex interactions between the blue force, the terrorists and the local population within which the actions are taking place in order to provide useful information to blue force enabling detection, neutralising or avoidance of IEDs.

Understanding of the numerous factors effecting IED use by terrorists can only be usefully developed by understanding the small group dynamics that we now understand to drive terrorist behaviour (Bohorquez et al, 2009) within the larger context of Counter Insurgency operations. Although understanding the human dynamics is the key to success, the models must also take into account the technology available to both sides and the geospatial situation in which it is deployed.

While there is much overlap in the elements used in each type of red teaming application the requirements of computational red teaming are rather different for the two cases resulting from the difference in the operational tempo between them and the way in which the products are to be used. In particular the collection, management and interpretation of real time data present a challenge.

2. Computational IED Red Teaming Elements

Computational red teaming for counter IED operations is both a behavioural and a technical problem (Whitney et al, 2009). The technical part of the problem is well understood and is

discussed in section 2.2. The behavioural part of the problem is less well managed but has been subject to huge research efforts over the last decade resulting in much progress.

2.1 Behavioural Aspects

Underpinning this progress has been:

- a. The acceptance and understanding that terrorism today is not necessarily the result of poverty, poor services and lack of education, but rather as Sageman details, (Sageman 2008) it seems to result quite often from feelings about the suffering of others, vicarious poverty for example, and so finds many of its recruits from the middle class, often people with a good education. Sageman also makes the point that many contemporary terrorist groups are “cliques” and as such all group profiles need to be local rather than national or global.
- b. That personal bias and operational information are localised in time and space, so that lessons learned in one place at one time are not necessarily applicable to other situations. (Kilcullen 2007), and
- c. Progress in social network theory and particularly in the application of dynamic network theory and meta-network analysis. (Bohannon 2009, Carley et al 2007, Silverman 2007)

The general term “behavioural” is used here to include all those elements having to do with the effects of people, socio-economic, media responses and manipulation and politics. Of course human behaviour takes place within the context of the environment and so there is a range of interactions with the non-behavioural elements of the situation through the effects each has on the other in terms of place, logistics and other resource constraints and these will be discussed as they arise.

Based on this brief review the following list of top level behavioural components are likely to be required for any computational red teaming application:

- Network based modelling of the interactions of the terrorist cell that employs the IEDs with its socio-political, religious and economic environment. In the last several years the need for understanding these relationships has been crystallised by the employment of Human Terrain Modelling (Marr 2008, Silverman 2007) Social network modelling has been a key element in the research effort funded by the US Joint IED Defeat Organization (JIEDDO), particularly with work at CMU (Carley 2006, 2007, Dombroski, et al 2003, Silverman, et al White 2006)

- Various approaches are used within this broad category, including Bayesian behaviour networks (Whitney 2009) and performance modifying functions using Markov chain methods (Weaver and Silverman 2001) Social network theory has been adequately described by Carley (2007) based on a broad foundation of social and biological network theory. There has been some criticism that this approach to dealing with complex systems is not science because it is not testable or reproducible in the classic sense which is certainly true. However, this seems to miss the point of its application to red teaming for counter IED and counter insurgency work, where it is used as a decision support tool by the people who need decision support tools. If they continue to use it, as it seems they have, then it is useful across the spectrum of their operations. Sageman has disputed the claims made by McCulloh (Bohannon 2009) for the effectiveness of metanetwork based tools introduced into the Iraq theatre, and no doubt his points are valid to some degree. Yet the complex factors operating in Iraq make it difficult to discount that these tools have been of considerable use.

Another approach that has been employed is partial network overlay. This is a form of metanetwork analysis that considers the network characteristics associated with many different possible networks connecting terrorists and their activities, skills and resources, as well as social and cultural factors such as identity, propensity for supporting violence, propensity to act and support for a cause. (North and Macal 2005)

Kilcullen's presentation of September 2007 shows 19 groups or organizations that he thinks need to be considered in counter insurgency in Iraq. Of these at a minimum probably 7 need to be included in any computational model to support red teaming. These are:

- Blue – blue forces, national army, national police
- Red – terrorist cell, insurgent group
- White – international media and local media.

But this list leaves out the extremely important intelligence contribution that is made by contacts with ordinary citizens who may not belong to any of the organizations listed above, and whose allegiance may shift from time to time. These citizens often belong to numerous other organizations such as tribal groups, militias and political groups. Information from this type of source is addressed

through the elements found in the concept of Human Terrain Mapping (Marr et al, 2008) and similar approaches

Whitney et al (2009) in their paper on modelling IED attack developed a list of behaviours at the terrorist cell level, that directly relate to the planning, building and delivering of an IED attack. Because they are behaviours directly related to the execution of an IED attack, they are behaviours that are closely connected to the technical aspects of the operation. This list includes items such as:

- plan attack, obtain finances, obtain materials, recruit, train,
- build IED, cache IED, prepare site, and
- issue order to employ, move and place the IED, initiate attack, escape

Silverman (2010) provides a relatively exhaustive list of the types of data required to support “in-country” experts when describing his web based interview instrument FAIREST. The acronym stands for “Factors, Agents, Institutions, Resources, Economy, Supra-Systems (Politics, propaganda, etc) and Time and so covers just about all the information that is available about a society. The Human terrain mapping effort as described by Marr et al (2008) covers much the same ground, although the goal there is to know and understand the local population in order to win them to the counter insurgency effort rather than to gather intelligence directly for the support of red teaming.

The widely reported work of Krebs (2002) in reconstructing the network of the 911 terrorist group from open source material only, illustrates that although intelligence gathering will always be difficult, the application of suitable context informing tools such as network analysis can provide relatively robust information.

2.2 Technical Aspects

The technical aspects that a computational IED red teaming tool needs to include are as follows:

- the various technologies for use of explosives, triggers and other components and their current availability,
- the technology currently favoured by the terrorists in the locality of interest,

- a detailed locality map, showing streets and roads, buildings, vantage points, choke points, important government and religious and cultural locations, possible IED assembly and target points and egress routes,
- blue force counter IED technology available, such as vehicle types, electronic counter measures, physical counter measures, and detectors,
- blue mission descriptions to support randomizing of patrol and other mission route selection and activities.

2.3 Operational Aspects

Although likely IED placement sites are important, Kilcullen (2007) makes the point that of all the key locations related to an IED event, placement site is the least important. More important from the point of view of combating IEDs is to understand and model signs of early warnings of operations, such as observers or mobile phone traffic, identifying possible firing and assembly points and any signs of their occupation. It is also important to be able to trace back from assembly points to local and district caches and to have a representation of these based on understanding of local networks. Since it can be assumed that a set of SOPs similar to this will be in use by the blue force, the computational model for red teaming must have a representation of this chain of operational stages. Whitney et al (2009) in their development of a Bayesian net model identify a set of observables that are very similar to Kilcullen's suggestions.

3.0 Modelling Methods

The literature survey revealed a number of different modelling approaches or methods that have been proposed and used for counter insurgency applications, many of which can be employed for counter IED computational modelling. However, there is very little published on actual counter IED applications and even less on red teaming for counter IED. No doubt this is because

“It is likely that much of the fundamental research for the design of modeling methods and tools can be done in open venues by researchers with low or no security clearances, but any work that includes specific and current field information on individuals or groups, specifics

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

on friendly or adversary force capabilities, or detailed operational plans must of necessity be highly classified to prevent the adversary anticipation, adaptation, and/or exploitation discussed above.” (Zacharias, MacMillan and Van Hemel, Eds. 2008)

It is clear, however, from the always numerous advertisements on Google and other sources of positions available in the counter IED red teaming field, that much activity is ongoing.

This section discusses a number of the different methods that seem likely to be useful in a computational red teaming context, with the expectation that the complex and heterogeneous nature of the problem area means that no one method is likely to be suitable for use with all aspects, and nor should there be any expectation that such could be the case.

In addition it is unlikely that any model that is used to support red teaming will be able to be validated in the accepted manner of engineering or physical science practice. There are a range of less rigorous techniques used in the soft sciences that provide some level of validation, for example face validation by SMEs, but to a great extent in the case of computational red teaming, validation is likely to be judged by the process of being used in the field. If it is judged by experience to be useful then it can be taken as having some validity for that time and place. Checkland puts it quite succinctly:

“The method of conducting investigations in natural science, based on the three principles of reductionism, repeatability and the refutation of hypotheses (Checkland 1981, Chapter 2), has been so widely successful that it has become the dominating model of research activity. Many people in many different fields make the unquestioned assumption that ‘research’ means testing of hypotheses.

However, the hypotheses which natural scientists test concern the natural regularities of the universe, and all the evidence is that these are invariant: the inverse square law of magnetism is always that, every time it is tested no matter who (competently) does the testing. When we turn to human affairs, however, and to social phenomena, it is far from obvious that the same experimental hypothesis-testing approach applies.” (Checkland 1998, pp18-19)

The recent US national Research Council Report Behavioral modeling and Simulation (Zacharias et al, 2008) makes a very strong case for the use of behavioural modelling for military applications in general, and counter terrorist applications in particular, despite the many difficulties in data collection and verification and validation.

Because computational red teaming must deal with a complex and heterogeneous environment for which a number of different modelling techniques may need to be used a model architecture that supports the ready use of components using such different techniques is appropriate. This suggests an Agent Based modelling approach, but other purpose built systems are also possible. Another desirable attribute of the tool to be used for computational red teaming is to have it driven by a time ordered event queue, rather than using a time step based approach. This is not only because there are numerous “actors” involved in any IED event or because social network behaviour is so important. It also arises because the model architectures that can be developed within the agent based approach can be structured to support the use of many different modelling methods within the one modelling system, allowing the selection of the most appropriate modelling approach depending on the type of agent, the current role of the agent at any time in the simulation and the context that the agent is operating in. Achieving this with any single modelling method is usually impossible. (Bonabeau 2002, North and Macal 2007, Miller and Page 2007, Zacharias et al 2008) As discussed below the concepts attached to agent based “distillations” are unlikely to provide an adequate basis for effective computational red teaming applications.

The modelling methods identified in this review have been developed and used in a number of different scientific, business and military areas, but derive primarily from behavioural science. In particular, network models of various types are of major importance for the reasons outlined in section 2 – that is, it is of overriding importance to understand and model the social networks that describe the terrorist behaviour and the blue force SOPs if red teaming models are to be useful.

A series of papers in the publication Connections between 2003 and 2009 describe a number of network concepts that apply to social network and tools for social network analysis (SNA), including analysis of SN statistical properties (Huisman and van Duijn, 2003), visualization of complexity in networks (McGrath, et al, 2003), measures for linking two networks together (Flom, et al 2004), whether or not to include ego into personal network models (McCarty and Witich, 2005), the use of the NetSAS tool to investigate in-degree

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

centralization (Johnston, et al 2007), measurement of Tie strength in social networks (Petroczi, et al 2007), time changes in network connectivity from a Bayesian perspective (Adams, et al 2008), the propensity for actors in a network to aggregate around a few states regardless of network structure (Stocker, 2009) and discovery of covert nodes within a network (Maeno, 2009)

Carley (2007) provides a list of the challenges for the use of networks in dynamic socio-cultural applications from the perspective of the CMU CASOS group including geospatial and other cross- discipline linkages.

The effect of missing data in social networks on the inflation of measurement error are investigated by Kossinets (2008) who also finds that social networks with multiple interaction contexts have interesting properties due to the presence of overlapping cliques in single mode networks.

Horne and Myer (2004) describe the Data Farming iterative modelling method of gaining understanding of complex situations that has potential to deliver useful results to operational understanding, although it is hard to see how the distillation of agent based models that they suggest will lead to future improvements in the computational red teaming field. For example according to expert practitioners and researchers in the field of human behavior modeling (Silverman et al., 2002; Pew and Mavor, 1998), “a common central challenge now confronting designers of HBM (human-behavior-modelling) applications is to increase the realism of the synthetic agents' behavior and coping abilities. It is well accepted in the HBM community that cognitively detailed, "thick" models are required to provide realism. These models require that synthetic agents be endowed with cognition and personality, physiology, and emotive components. To make these models work, one must find ways to integrate scientific know-how from many disciplines, and to integrate concepts and insights from hitherto fragmented and partial models from the social sciences, particularly from psychology, cultural studies, and political science. One consequence of this kind of integration of multiple and heterogeneous concepts and models is that we frequently end up with a large feature space of parameters that then need to be filled in with data.” (Silverman, et al 2008)

Yet another method that has been developed for counter terrorism modelling is that of analogical reasoning. Markman, et al (2003) discuss several analogical reasoning methods for use in intelligent

counter terrorism systems and show why it is superior to case base reasoning, which lacks the flexibility necessary for the dynamic situations that exist in trying to combat terrorism.

There are many different techniques that are used to model human behaviour, and many of these have been used to model insurgent behaviour and could potentially be used to develop elements in a computational IED red team application. The following is a list of some of the more interesting papers found that exemplify the application of several of these methods or that describe interesting extensions to the basic methods:

- Analysis of Competing Hypotheses
 - Valtorta, et al (undated) An application of this method to complex decision analysis
 - Pope, et al 2006 provide a discussion of weaknesses and an extension for use in intelligence collection
- Artificial Neural Networks
 - Schmidhuber, J. 1993 Describes an ANN that can reason on and modify its own weights
 - Marchiori, D. and Warglien, M. Predicting Human Interactive Learning by Regret-Driven Neural Networks. 2007
- Bayesian approaches
 - Sentz and Ferson 2002 Dempster – Schafer theory of belief functions is a generalization of Bayesian theory for probabilistic uncertainty
 - Brooks, S. 2004 Markov Chain Monte Carlo method that assists in obtaining posterior distributions
 - Whitney, et al 2009
 - Adams, et al 2008 used a Bayesian method to investigate time changes in connectivity in a social network. The code Winbugs was used to simulate the posterior distributions.
- Genetic Algorithms
 - Mitchell, et al 1992 which has an excellent discussion on fitness landscapes
 - Parunak, et al 2007 use a GA when empirical data is available to configure the personality vectors of the swarming ghost entities in their DEFUSE model for analysing IED threats.

- Game Theory
 - Kott and Ownby, 2005 discusses the game theory based approach adopted for the Adversarial Reasoning Module in the DARPA RAID program (Real Time Adversarial Intelligence and Decision Making)
- Performance Modifying Functions
 - Silverman, B. G. 2007
 - Silverman, B. G. , et al 2003
 - Weaver, et al 2001
- Systems dynamics
 - Anderson 2006
 - Brothers and Pavlov (undated) show a causal loop diagram and a stock and flow diagram for a systems dynamic model of IED effectiveness.
- There is a collection of work on the modelling of asymmetric adaptive adversaries. (eg Brims and Ludwig 2008, Ludwig and Farley 2008, Ponson and Spronk 2004) While this work is based mostly on games development there is useful insight that could drive new methods for developing models of terrorists engaged in IED activities.

The Defense Science Board Study “Understanding Human Dynamics”, March 2009 contains a very large amount of information about the needs and on-going work in the US. There is a gap analysis table on page 52 which lists the need for a number of computational R&D programs , some of which will impact on any computational red teaming development efforts as they focus on behavioural, economic, cultural and perception based issues, and in particular the need for cross boundary network analysis. There is a list of over 20 major US Defense organisations that have active programs that bear on human dynamics beginning on page 100, many with multiple programs, and on page 54 there is a list of DARPA programs. The report makes the point that the social net work analysis tools are relatively mature and do not need to be re-invented, but that human dynamics computational modelling is not mature. The report also highlights the data problem that all such modelling must develop methods to address.

3. Data Collection and Processing Tools

This discussion on data collection and analysis tools is confined to those tools related to specific human behaviour modelling approaches found in the counter insurgency literature, and so does not

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

attempt to cover the huge literature on general data collection and analysis methods . It also does not attempt to cover specific intelligence organisation approaches to data collection and analysis.

Social network theory and network analysis can provide a strong contextual framework for the collection of the type of data required to support computational red teaming. Carley (2007) in a CASOS presentation describes the capability of the tool AutoMap which uses a combination of statistical and positional information to generate possible network information by scanning to discover patterns and vulnerabilities. (ORA is an open source environment, but AutoMap, while downloadable in script format, is not) Carley's group has also developed a Dynamic Network Analysis (DNA) kit to work with the ORA suite. (www.casos.cs.cmu.edu/projects/automap/software.html)

The US Army Human Terrain Mapping approach collects similar data about the local environment and collates it in the Command Post of the Future computer systems, from whence it can be shared across local units and up the command chain. This is tactical data for use on a daily basis and therefore is suitable for use in operational red teaming computational modelling. Silverman et al (2007) have extended their PMFServ modelling system to use this type of information and suggest further extension to analysing alternative competing hypotheses for DIME-PMESII studies.

Stafford (2009) used the CARMA tool (Hidber 1999) to develop association rules to model commonly occurring sequences of events leading to IED attacks. CARMA puts association rule mining online, in that it gives continuous feedback, is user controllable and yields deterministic and accurate results.

stOCNET is a very useful open source tool for the statistical analysis of social networks that provided tools for most of the newer network statistical measures that might be needed to extract useful data for red teaming modelling (Huisman and van Duijn, 2003)

Priebe et al (2005) used a large data base of ENRON emails to demonstrate the use of scan statistics in detecting anomalies in email traffic in relation to ongoing crisis events inside the company. It seems likely that similar methods could be used to gather and analyse information about unusual terrorist activity levels in the context of larger scale developments across an AOR. Whether this could work on a local scale for operational red teaming is an open question.

Pope, et al (2006) discuss the weaknesses of the ACH approach in regard to understanding deception and misperception errors in intelligence collection which provides some useful insight into

the use of subjective logic within the ACH framework help considerations of the type of data collection processes required to support computational red teaming.

There are many other data modelling, analysis and collection methods in the literature, including for example, the modelling approach Analysis of Competing Hypotheses (ACH) which can provide a formal and useful method for assessing the validity of intelligence data (Pope et al 2006), machine statistical learning (Ding, et al 2009) the generation of “data” for use in modelling computer agents (Ludwig and Farley, 2008) and artificial markets (Zenobia, 2008), but time and resource have not allowed for these to be researched for this paper.

Backus and Glass (2006) provide an excellent analysis of the factors required to model terrorist behaviour using an agent based context. In particular their discussion of agent coping skills in relation to responding to challenges is worthwhile.

4. Toolkits and Computational Environments

There are several important examples of toolkits and descriptions of development processes for counter insurgency application that translate well to an IED red teaming application. The first of these is the large body of work done at the Center for Computational Analysis of Social and Organizational Systems (CASOS) at CMU under Kathleen Carley. A particularly useful paper is that by Carley et al (2007a) describes the assembly of a tool set that includes automated data collection and processing to provide an empirical basis for an initial network structure and the use of a meta-matrix model to manage ontology issues, use of the open source tool ORA for networks analysis, simulation of agent interactions across the networks using an agent based model system DyNet leading to re-analysis of the network structure with ORA. The paper contains a useful worked example and discussion of the tool AutoMap that is used to extract the empirical data from a set of texts. This paper describes some of the CASOS initial ideas on meta-network analysis for counter terrorism work described in Bohannon (2009) and elsewhere by other CASOS papers. (Carley 2007b) The advantages of combining network model analysis with multi-agents simulations to address the possible behaviours of the agents in a network are: (Carley 2007a) “There are a number of advantages to a statistical network analysis tool like ORA that originate from implementing the meta-matrix. The possible analyses are more comprehensive and provide greater insight into factors that drive behavior in comparison to tools that are restricted to operate on the agent level. The types of analyses supported with ORA include:

- “ Identification of weak and strong agents or organizations in a network, points of influence, hidden substructure, organization's capabilities.
- Optimization of an organization's structure for various outcomes including general high performance and Adaptivity.
- Comparison of an organization with other organizations, random networks of equivalent size, or the same organization after an intervention or at a different time point.
- Identification of the sphere of influence surrounding specific agents or organizations.”

Another group that has done a great deal of counter insurgency computational tool building and modelling is Silverman et al at Pennsylvania University. This work is based on a number of year's development of the Performance Modifying Functions (PMF) paradigm, which also involves a great deal of empirical data in the production of the toll PMF Serve. Unlike the CMU tools this is a proprietary product, but it has found considerable use within the counter terrorism modelling efforts supported by DARPA and other US Agencies and is discussed further in Section 6. Weaver, et al (2001), Silverman et al (2002), Silverman et al (2003), Silverman (2007).

The DIME PMESII community, supported by DARPA and the US Services has also produced many useful tools for understanding and modelling the complex socio-political –economic interactions involved in counter terrorism and asymmetric warfare. A typical example is the Toolkit for Building Hybrid, Multi-resolution PMSEII Models, from Charles River Analytics, Inc. This work used the proprietary tool GRADE (Graphical Agent Development Environment) for the development of a component based architecture that enables the use of different modelling paradigms such as semantic models, network models, systems dynamic models, existing PMESII models and graphical display components into the same architecture. (Bachman & Harper, 2007) The capability to have this type of architecture available to support computational red teaming seems essential.

A useful discussion on the creation of DIME/PMESII models is provided by Partnow and Hartley (2002). Their structured top down and bottom up approach explores the linking of anthropological knowledge and OR techniques to create the useful models.

Some of the literature suggests that games technology can be useful, while others (Ref here) doubt that it provides the necessary rigor for successful use. It would seem that some techniques in use for game building would be very useful, as for example the dynamic scripting tool for developing Asymmetric adversaries described by Ludwig and Farley (2007)

van Dyke Parunak, et al (2006) describe a model for introducing emotion into the behaviour of agents in a combat situation that extends the treatments used in some earlier models such as EINSTEIN (Ilachinski, 2004) and MANA (Lauren and Stephen, 2002) with some much more realistic representations. This model and the analysis of requirements that it presents provides a useful introduction to this aspect of modelling agent behaviour should it prove necessary to have emotion included in the computational red teaming applications.

In addition to the purpose built toolkits discussed above there are a number of robust general agent based toolkits that are commonly used to develop models in this domain. The most prominent of these, all of which are open source and can be downloaded from the web, are probably:

ASCAPE (<http://ascape.sourceforge.net>)

Mason (<http://cs.gnu.edu/~eclab/projects/mason>)

Repast (<http://repast.sourceforge.net>)

Swarm (http://swarm.org/index.php/Swarm_main_page)

Berryman (2008) provides a useful review of 8 agent based toolkits.

5. Specific Computational Models

The open literature contains very little on specific red teaming tools, computational or otherwise. (See Section 3 or Zacharias et al 2008) There are, however, a handful of papers that deal specifically with the IED problem from a more general view point each of which makes useful contributions to the concepts that a computational red teaming application would need. In addition there is a rich literature discussing a variety of computational models using many different methods for addressing general counter terrorist or counter insurgency applications. Much of this work could also find direct application to building and using a computational red teaming environment and this section attempts to provide insight into some of these connections.

The Center for Non-Proliferation Studies (2002) prepared an extensive literature review of terrorist behaviour modelling for DTRA that provide a long list of practitioners together with a rating scale from 1 to 5 indicating the relevance to the topic that is a useful source – although none of the work is related to IED red teaming.

The DEFUSE model of Parunak, Sauter and Crossman (2007) is a 3 layer hybrid agent based model that builds on work for the DARPA RAID program (2004) by using the Adversarial Reasoning Module to reason about IED locations, frequency and tactics, the Power Struggle Toolkit (PSTK) under the DARPA COMPOEX program to build a network model of faction leadership and a swarming agent on a network concept from the ONR CIED program (2009) to develop many possible futures which it then uses to prepare probability distributions and event forecasts for likely IED events. The leadership models are developed by SMEs for incorporation into PSTK and in practice this type of understanding could be developed from Human Terrain data. Its hybrid architecture “applying the structure and technology that best solves the problems presented in each layer” seems to be an excellent in principle approach to the type of model required for a red teaming IED tool.

Silverman (2007) provides another hybrid approach to modelling IED events and blue response to them. The paper “Human Terrain Data – What should we do with it” uses the Performance Modifying Function approach (using the COTS product PMFServ) as a cognitive model for simulating agent behaviour. PMF serve has found application in a number of other areas and is a mature product as the result of a very large development effort at Pennsylvania University over the last decade. FactionSim, a module that allows simulation of inter-group competition, is then used to “embed PMFServ within a Game theory/PMESII campaign framework”. In its concluding remarks this paper also makes useful observations about the testing and correctness for this type of modelling.

The work reported from the ONR CIED STIFLE project (2009) is also of interest. The project had three major objectives – “to extend the predictive polyagent modeling construct to include explicit reasoning over task execution by individuals and groups, to develop theoretical, formal and experimental analysis tools and methods to characterize and influence the dynamics of predictive polyagent models and to apply the extended modeling and analysis capabilities to the problem of IED prediction and forensics.” Of particular interest is the idea of deploying swarms of fine-grained agents that move probabilistically on a multi-pheromone landscape. Although this stigmergic process is a concept that is not very realistic in that it produces agent behaviour without any direct action between the agents, it provides a good method to model the effects of numerous spatially and/or temporally distributed influences on agent behaviour when specific individual agent information is lacking.

Whitney et al (2009) have built a model that enables forensics and prediction based on what they call an IED kill chain model. This useful work identifies observables at various stages along the net that can be used to detect the likelihood of an attack.

Another approach suggested by Stafford, in a Masters thesis at NPGS (2000) uses the CARMA algorithm (Hidber 1999) to develop a sequential pattern detection model that together with time series analysis simulates the timing and frequency of IED attacks.

Models of Insurgency and for Counter-Insurgency application are rather more common than those with a specific IED focus, and several are mentioned here to illustrate the range of methods that might well have application to computational red teaming.

Anderson (2006) used a systems dynamics method to model the case of the Anglo-Irish war of 1916-1921, and provides a useful discussion of the strengths and weaknesses of the method. His conclusion is that agent based modelling is probably a better approach under the conditions usually found in insurgency situations.

An agent-based model has been used to develop a counter IED (CIED) red-teaming model at DSTO JOD. The goals of this agent-based model are: to identify possible trends in IED usage, and to gauge these with respect to various countermeasures. The initial development of the model was to allow exploratory, “what-if?” type questions to be asked, but with the possibility of being used as a predictive tool after extensive ongoing validation of model performance against observations.

The agent-based model has three types of agents:

Blue agents, which conduct operations along user-defined routes, which employ a range of different vehicles and countermeasures.

White agents which represent the general populace that provide information to both blue agents (about IEDs that have been placed) and red agents (about routes of blue convoys so that they may be attacked in the future).

Red agents which represent insurgent / terrorist cells that undertake and/or support and direct IED operations against blue and white agents. Red agents learn about combinations of IED components and have a more complex implementation allowing goal selection modified by perceived experience.

A Petri Net is used for distributing red objective information as well as available resources such as funding and supplies. The actual network configuration can be changed to reflect changing knowledge about red agents in theatre. External events, such as the introduction of new components, new counter measures, changes of blue routes and red target selection in response to perceived levels of past success can be introduced at various time points through the simulation.

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

In response to the general perception that it is essential to engage the local populations in all phases of counter insurgency operations including countering IEDs it is planned that a more detailed model of “white” agents will be developed for integration with the existing CIED model.

An agent based model developed by Tsvetovat and Carley (2002) allows simulation experiments to test the impact of different isolation strategies and wiretap strategies on performance of the terrorist organization. A number of intelligent strategies are compared to a baseline strategy of isolating random agents. The model uses sets of apposing terrors and police agents, each with imperfect knowledge about the others organization.

Argonne National Laboratory developed a multi-layered network agent based model for DTRA called Netbreaker (North and Macal 2005)

- NetBreaker uses agent-based social modeling to find possible terrorist networks bounded by:
 - Known computable rules of social network formation
 - A given list of participants, along with possible unknown players
 - Existing evidence documenting interactions between the participants, along with possible unobserved, but hypothesized, interactions
- The result is a “space” of possible terrorist networks:
 - If the list is large enough, then the space of alternatives will contain the actual network being investigated
 - This space of alternatives can be used to create actionable questions that narrow the possibilities for the actual network
- NetBreaker’s design goal is to reduce surprise by providing and quantifying possibilities, not to determine by itself which possibility is correct
- NetBreaker does not remove human analysts from the investigative process, but instead helps them consider more possibilities than they could have before.

The CASOS group series of “Challenge” versions of the BioWars model is not directly relevant to red teaming, but offers many ideas on how to assemble an agent based model that includes many different and heterogeneous components, such as weather and climate, disease vectors, geospatial information, first response organization, information flows and societal structure. (Carley, et al 2004)

The NonKin Village game developed by Silverman (2009) develops a method to integrate cognitive agents, social agents, economic models, and story generation tools to construct a cultural-based human terrain in which there are a number of different tribes with differing histories in a single village. Although the objective of this model is training, it also contains many useful ideas on how to build a complex agent based model.

DSB Report, (2009) Appendix E gives a useful summary of the large number of computational modelling tools and methods that are now available.

6. Summary and Conclusions

This survey of the literature relating to computational red teaming has adopted the view that the key element of any such tool will be the use of social network representations of the behaviour of the terrorists and of the local population, which may or may not support the use of IEDs.

The paper begins with a discussion of the computational elements that are likely to be required for a computational red teaming application, consisting of representations of the behavioural, technical and operational aspects of the domain. This is followed by discussion of the most often used methods for modelling these components, including some useful specific application examples. Next there is a section on data collection and processing tools that are used. This is an important consideration in any red teaming exercise because most of the models used will be data hungry. The next section presents some of the major toolkits and computational environments that are reported in the literature, with a heavy concentration on dynamic social network tools. The final section presents brief discussions of a number of actual models, some of which are specifically addressed to the IED problem, but most of which are more generally counter insurgency models that are likely to contain useful ideas for developing a computational IED application.

Most of the sources quoted were found on-line, and are nearly all US in origin, which is a reflection of the huge amount of work that has been funded in the US since 2001. Also, nearly all the research that is reported relates to counter terrorism and counter insurgency – there is relatively very little on real computational red teaming applications and even less

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

on red teaming for counter IED application, despite there being plenty of evidence that a great deal of work is being funded. There is also the view among experienced red team practitioners that computation has a very limited role to play, but this seems to arise from a misunderstanding – computational red teaming will always be a support tool to analysts, to expand the domain that they can consider and to allow their exploration of complexity.

There is an enormous amount of information about basic social network science, its exploitation by computer models for counter insurgency applications, on-going research efforts, including funding agencies and research teams and defense (US) guidance contained in just the following six sources – DSB 2008, Zacharias, et al 2008, Jackson 2008, FM 3-24, Center for Non-Proliferation Studies 2002 and the CASOS web site, www.casos.cs.cmu.edu.

There is a distinction to be made between an operational red teaming application and one that is to be used for forward planning and capability assessment, driven by the difference in time factor which reflects immediately on the amount and type of information that can be gathered and analysed.

The most recent literature all indicates that understanding the human environment is the key to successful management of the IED problem. Any computational tool will need to be supported by effective data collection and analysis tools that are specifically structured to support the modelling needs. This need has been recognized by the Human Terrain program initiated by the US Army, and several other similar initiatives. The same sources make the point that the key to effective counter IED action lies in working upstream from the actual emplacement of the IED, so the red team must be able to represent this part of the environment accurately which will require accurate geospatial data, including buildings, roads, meeting points and other local points of interest together with an understanding of the movement of people and material resources through this landscape.

Modelling this multi-dimensional and heterogeneous domain is probably best done using an agent based architecture, and most of the examples in the literature have used this to some degree, if not completely. This means agents in the broadest sense of the word – that is, any entity that has a state that might change, for any reason. Thus passive agents are a useful concept within the overall architecture to make it easier to employ the different types of

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

modelling techniques required to best represent the various activities that need to be represented. Two other design principles that are likely to lead to more effective modelling are the use of a time ordered event queue and an architecture that supports the separation of what is being modelled from how it is being modelled. This latter provides for greater flexibility and extensibility as more information becomes available. The literature also points towards the development of “thicker” agents and away from agent based distillations which use simplified or unrealistic processes to gain realistic results.

The disadvantages in this approach are that validation in the “hard” science and engineering sense is usually not possible, and that the models are data hungry. Research on such ideas as performance modifying functions at the University of Pennsylvania and the CASOS group work using huge amounts of open source data may supply some basic behaviour understanding in a useful domain setting that has been largely lacking in the more general social behaviour literature. Advanced risk management techniques will be needed to ensure that data collected from the field is reliable, and several approaches to satisfying this need were found.

There has been some trenchant criticism of both the use of such models in Iraq and Afghanistan, and of some of the claims made regarding their usefulness, but involvement with and adoption by the end user stakeholders will settle this question over time. The US Defence Science board has strongly supported further research and development and any work of this type that leads to greater understanding of the dynamics around the use of IEDs would seem to be beneficial.

References

Adams, S., Carter, N., Hadlock, C. and Sirbu, G. Change in Connectivity in a Social Network over Time: A Bayesian Perspective Connections 28(1) 2008

Anderson, E. G. Modeling Insurgencies using Systems Dynamics: The Anglo-Irish War, 1916-1921 as a Case Study, 2006. University of Texas, at www.EdAnderson.org

Bachman, J. A and Harper, K. A. A Toolkit for Building Hybrid, Multi-resolution PMSEII Models, AFRL-RI-RS-TR-2007-238 Final Technical Report November 2007, Air Force Research Laboratory Information Directorate, Rome, New York

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

Backus, G. A. and Glass, R. J. An Agent-Based Model Component to a Framework for the Analysis of Terrorist-Group Dynamics SANDIA REPORT SAND2006-0860P Unlimited Release February 2006

Berryman, M. Review of Software Platforms for Agent Based Models. 2008 DSTO-GD-0532
Land Operations Division Defence Science and Technology Organisation

Bohannon, J., Counterterrorism's New Tool: Metanetwork Analysis. Science V325. 2009

Bohorquez, J. C., Gourley, S., Dixon, A. R., Spagat, M. and Neil F. Johnson, N. F. Common ecology quantifies human insurgency, Nature 462, Dec. 2009

Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems PNAS May 14, 2002 vol. 99 suppl. 3

Brooks, S. Markov Chain Monte Carlo and Gibbs Sampling 2004
MCMC Preprint Service: <http://www.maths.surrey.ac.uk/personal/st/S.Brooks/MCMC/>

Brothers, A. and Pavlov, O. V. Using System Dynamics to Model Risk Perception and Communication in Response to Threat. (Undated)<http://www.decisionresearch.org/pdf/BrothersPavlov.pdf>

Carley, K. M., Altman, N., Kaminsky, B., Nave, D. and Yahja, A. BioWar: A City-Scale Multi-Agent Network Model of Weaponized Biological Attacks. , 2004

Carley, K. M., "Dynamic Network Analysis" in the Summary of the NRC workshop on Social Network Modeling and Analysis, Ron Breiger and Kathleen M. Carley (Eds.), National Research Council. 2006

Carley, K. M., Diesner, J., Reminga, J. and Tsvetovat, M. Toward an interoperable dynamic network analysis toolkit. Decision Support Systems 43 (2007) 1324–1347

Carley, K.M. Network Science: The Dynamic Socio-Cultural Perspective. CASOS PowerPoint presentation, 2007 CASOS, ISR, SCS

Center for Non-Proliferation Studies. Literature Review of Existing Terrorist Behavior Modeling, for Defence Threat Reduction Agency, August 2002

Checkland, P., and Holwell, S. (1998). Information, Systems, and Information Systems: Making Sense of the Field. Chichester, Sussex; New York: Wiley.

Counter-Insurgency. FM 3024 (MCWP 3-3.5) Department of the Army. Dec 2006

Defense Science Board, Defense Science Board Task Force on the Role and Status of DOD Red Teaming Activities, Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Washington, D.C., September 2003

Defense Science Board, Defense Science Board Task Force on Understanding Human Dynamics. Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, Washington, D.C., March 2009

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

G. Ding, R., Lax, J., Chen, P., Chen, and Marx, B. Learning Terrorist Profiles by Optimization with Two Objectives. World Comp 2009

Dombroski, M. , Fischbeck, P. and Carley, K. M. Estimating the Shape of Covert Networks Proceedings of the 8th International Command and Control Research and Technology Symposium, Washington, DC 2003

Flom, P. L., Friedman, S. R., Strauss, S. and Neagius, A. A New Measure of Linkage Between Two Sub-networks. Connections 26(1) 2004

Hidber, C. Online Association Rule Mining. In Proceedings of the ACM SIGMOD International Conference on Management of Data, 1999

Horne, G. E. and Meyer, T. E. Data Farming: Discovering Surprise Proceedings of the 2004 Winter Simulation Conference, R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

Huisman, M. and van Duijn, M. A. J. StOCNET: Software for the statistical analysis of social networks. Connections 25(1) 2003

Ilachinski, A. Artificial War: Multiagent-based Simulation of Combat. Singapore, World Scientific, 2004

Jackson, M. O. Social and Economic Networks, Princeton University Press 2008

Johnston, M. F., Chen, X., Bonacich, P. and Swigert, S. Modeling Indegree Centralization in NetSAS: A SAS Macro Enabling Exponential Random Graph Models. Connections 27(2) 2007

Kilcullen, D. Counter Insurgency in Iraq: Theory and Practice, 2007. Noetic

Kott, A. and Ownby, M. Decision Making and Cognitive analysis. 2005 Defense Advanced Research Projects Agency

Kossinets, G. Effects of missing data in social networks. 2008 Institute for Social and Economic Research and Policy, Columbia University, 420 W. 118th St. email: gk297@columbia.edu.

Krebs, V. E. Mapping Networks of Terrorist Cells, Connections 24(3): 43-52
2002

M. K. Lauren and R. T. Stephen. Map-Aware Non-uniform Automata (MANA)—A New Zealand Approach to Scenario Modelling. Journal of Battlefield Technology, March 2002

Ludwig, J. and Arthur Farley, A., Using Hierarchical Dynamic Scripting to Create Adaptive Adversaries 2007. Stottler Henke Associates San Mateo, CA 94404

Malone, T. G. and Schaupp, R.E. The Red Team: Forging a Well-Conceived Contingency Plan. Aerospace Power Journal, Summer 2002

Marchiori, D. and Warglien, M. Predicting Human Interactive Learning by Regret Driven Neural Networks. 2007 Advanced School of Economics and Department of Business Economics and Management, Ca' Foscari University, Venezia, Italy.

Markman, A. B., Rachkovskij, D. A., Misuno, I. S. and Revunova, E. G. Analogical Reasoning Techniques in Intelligent Counterterrorism Systems. International Journal "Information Theories & Applications" Vol.10 139, 2003

Marr, J., Cushing, J., Garner, B. and Thompson, R. Human Terrain Mapping: A Critical First Step to Winning the COIN Fight. Military Review. March-April 2008

McGrath, C. and Witich, A. Conceptual and Empirical Arguments for Including or Excluding Ego from Structural Analyses of Personal Networks. Connections 26(2) 2005

Miller, J. H. and Page, S. E. Complex adaptive systems: an introduction to computational models of social life. Princeton University Press. 2007

Mitchell, M., Forrest, S. and Holland, J. H. The Royal Road for genetic Algorithms: Fitness landscapes and GA Performance. In Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life Cambridge, MA: MIT Press, 1992.

North, M. and Macal, C. Netbreaker Terrorist Network Simulation. 2005. Center for Complex Adaptive Agent Systems Simulation (CAS2) Decision & Information Sciences Division, Argonne National Laboratory

North, M. and Macal, C., Managing Business Complexity: Discovering Strategic Solutions with Agent Based Modeling and Simulation Oxford University Press. 2007

van Parunak, H., Bisson, R., Brueckner, S., Matthews, R., John Sauter, J. A Model of Emotions for Situated Agents. 2006 Altarum Institute, 3520 Green Court, Suite 300, Ann Arbor, MI 48105-1579 USA

van Parunak, H., Sauter, J. and Crossman, J. Multi-Layer Simulation for Analyzing IED Threats, 2007 <http://www.newvectors.net/staff/parunakv/HST09MultiLayerSim.pdf>

Partnow, P. H. and Hartley, D. S. Using Cultural Information to Model DIME/PMESII Effects. 2002. Hartley Consulting

Petroczi, A., Nepusz, T. And Bazso, F. Measuring tie-strength in virtual social networks. Connections 27(2) 2007

Pew, R. W. and Mavor, A. S. (Eds) Modeling Human and Organizational Behavior: Applications to Military Simulations. 1998 National Research Council

Priebe, C. E., Conroy, J. M., Marchette, D. J. and Park, Y. Scan Statistics on Enron Graphs, SIAM International Conference on Data Mining, Workshop on Link Analysis, Counterterrorism and Security, Newport Beach, California, April 23, 2005.

Pope, S., Josang, A. and McAnally, D. Formal Methods for Countering Deception and Misperception in Intelligence Analysis COALITION COMMAND AND CONTROL IN THE NETWORKED ERA 2006, 11th. IICTRS

REAL-TIME ADVERSARIAL INTELLIGENCE AND DECISION-MAKING (RAID), DARPA SOL BAA 04-16, 04/21/04

Sageman, M. Understanding Terrorist Networks. University of Pennsylvania Press 2004

Sageman, M. Leaderless Jihad, University of Pennsylvania Press 2008

Schmidhuber, J. A Neural Network that Embeds its Own Meta-Levels. Proc International Conference on Neural Networks. San Francisco IEEE 1993

Sentz, K. and Ferson, S. Combination of Evidence in Dempster-Shafer Theory. T.J. Watson School of Engineering and Applied Sciences, Binghamton University, NY 2002

Silverman, B. G., Johns, M. Shin, H. and Weaver, R., Performance Moderator Functions for Human Behavior Modeling in Military Simulations, 2002, Prepared for the Human Behavior Program, Defense Modeling and Simulation Office

Silverman, B. G., Cornwell, J.B. and O'Brien, K. Progress to Date on the Human Performance Moderator Function Server (PMFServ) for Rapidly Generating Reusable Agents, Forces, and Crowds, 2003, Ackoff Center for the Advancement of Systems Approaches (ACASA) University of Pennsylvania, Towne 229c Philadelphia, PA 19104-6315

Silverman, B. G. Human terrain Data – What should we Do With It? Proceedings of the 2007 Winter Simulation Conference, S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds.

Silverman, B. G., Bharathy, G. K. and Kim, G. J. Challenges of Country Modeling with Databases, Newsfeeds, and Expert Surveys. 2008 ESE Papers, University of Pennsylvania

Silverman, B. G. Systems Social Science: A Design Inquiry Approach for Stabilization and Reconstruction of Social Systems, Intelligent Decision Technologies V4 (1), 2010

Stafford, W. B. Sequential Pattern Detection and Time Series Models for predicting IED Attacks. 2009. Thesis -US Navy Post Graduate School, Monterey.

Tsvetovat, M. and Carley, Kathleen. (2002). Knowing the Enemy: A Simulation of Terrorist Organizations and Counter-Terrorism Strategies. CASOS Conference 2002, Day 3, Electronic Publication, Pittsburgh, PA.,
webSite:<http://www.casos.cs.cmu.edu/events/conferences/2002/pdf/day2.pdf>

Valtorta, M., Dang, J., Goradia, H., Huang, J. and Huhns, M. Extending Heuer's Analysis of Competing Hypotheses Method to Support Complex Decision Analysis. Dept Computer Science & Engineering, University of South Carolina (undated)

Defence and Systems Institute, University of South Australia, F2-23 Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia P (08) 8302 3369 • F (08) 8302 5344 • www.unisa.edu.au/dasi

Weaver, R., Silverman, B. G., Shin, H. and Dubois, R. Modeling and Simulating Terrorist Decision-making: A 'Performance Moderator Function' Approach to Generating Virtual Opponents Proceedings of the 10th Conference On Computer Generated Forces and Behavioral Representation, May 2001.

White, J. B. The Challenge of Adversary Networks in Iraq. The Washington Institute for Near East Policy, March 2006

Whitney, P., Thompson, S., Brothers, A., Coles, G., Henderson, C., Young, J., Niesen, D. and Madsen, J. Techno-social Modeling of IED Threat Scenarios and Attacks. 2009 Pacific Northwest National Laboratory, PNL- SA- 65672

Zacharias, G. L. MacMillan, J. and Van Hemel, S. B. Eds. Behavioral Modeling and Simulation: From Individuals to Societies, 2008. National Academies Press

Zenobia, B., Webe, C. and Daim, T. Department of Engineering and Technology Management, Portland State University, Post Office Box 751, Portland, OR 97207-0751, USA Available online 31 October 2008

COMPUTATIONAL RED TEAMING – A LITERATURE SURVEY AND COMPUTATIONAL TOOL REVIEW

A REPORT FOR THE DEFENCE SCIENCE AND
TECHNOLOGY ORGANISATION'S JOINT
OPERATIONS DIVISION
APRIL 2010

Table of Contents

Introduction	2
Approach	2
Improvised Explosive Devices (IED's) – An Example Scenario.....	3
Background on Computational Red Teaming	6
Agent-based Modelling and Simulation	6
Agent-based Distillations	8
Optimisation methods for CRT	8
<i>Evolutionary Algorithms</i>	10
<i>Co-evolution</i>	10
<i>Swarm-based optimisers</i>	12
Data Farming	13
Literature survey on Computational Red Teaming.....	13
IEDs	16
Computational Tools for Red Teaming.....	16
<i>Isaac and Einstein</i>	16
<i>MANA</i>	17
<i>Pythagoras</i>	19
<i>Crocodile</i>	20
<i>Wisdom and WISDOM-II</i>	21
<i>PAX (and PAX3D)</i>	21
<i>General Purpose Multi-agent Simulation Tools</i>	22
<i>Prototype or Special Purpose Tools</i>	23
<i>A Comparison</i>	24
Tools used in Recent Data Farming Workshops	24
Conclusion and Future developments	25
References.....	26

COMPUTATIONAL RED TEAMING – A LITERATURE SURVEY AND COMPUTATIONAL TOOL REVIEW

A REPORT FOR THE DEFENCE SCIENCE AND TECHNOLOGY

ORGANISATION'S JOINT OPERATIONS DIVISION

INTRODUCTION

This report has been produced at the request of the Joint Operations Division of DSTO (Defence Science and Technology Organisation). The brief was to conduct a literature and computational tool review on computational red teaming, and to produce a report for presentation at a DSTO-sponsored workshop.

Red teaming is a method for assessing vulnerabilities in systems or structures. Two factions or teams – red and blue – are posited or formed. The red team is charged with attacking the system or structure being defended by the blue team. The role of the red team is to challenge the implicit assumptions in blue team defences. Although the idea originated in the context of military simulations or wargaming, it can be applied more broadly to, for example, civil defence scenarios, security assessment, business decision-making, realtime strategy games, and computer network vulnerability assessment.

Red teaming can also be used for training purposes – to allow participants to learn and be tested in a safe environment.

Traditionally, red teaming has been done manually, either in a manual simulation on a board or table, or in physical wargaming, or with real teams physically infiltrating a secure facility. More recently, computer-based simulations or computational models are also used - this is what we refer to as *Computational Red Teaming* or CRT.

Compared to manual red teaming, computational red teaming has the advantage that many scenarios can be simulated and analysed. This makes it possible to gather enough data to carry out statistical analyses, or to use search-based optimisation and machine learning methods. These methods can also often uncover unexpected knowledge, as computerised methods do not suffer from “blind spots” in the same way that humans do.

APPROACH

Because of our particular interest and expertise, in this report, we place most emphasis on the *Optimised Red Teaming* approach, which combines agent-based simulation with evolutionary optimisation methods. We also provide an overview of CRT more generally, placing optimisation methods within the

CRT context. As well as a literature review, we have also produced a video presentation illustrating the approach on an example IED scenario.

Two specific topics that were identified as of interest in the brief for this research are applications of CRT to *Improvised Explosive Devices (IEDs)*, and the use of *co-evolution* methods.

From Schedule 1 of the Research Agreement (**bold added**):

*“With DSTO obviously linked to Defence, the specific focus for this work is Computational Red Teaming (CRT) **as applied to Improvised Explosive Devices (IEDs)**. An initial (constrained) layer is to address a specific event – rather the:*

When: Triggering events, Time requirements

Where: City, Building, Road (specific geographic location), and

How: Approaches Required, Technology, Funding, Know-how.

...

*A broad array of linked topics have been identified for the literature and computational tool survey, and they range across artificial markets, statistical learning, data farming networks, **co-evolution methods**, cyber warfare, risk assessment, game theories, games with IED scenarios, agent based distillation, fitness landscape and adversarial learning.”*

The brief also foreshadows follow-on activities, including the use of evolutionary algorithms in conjunction with some computational tool (we would suggest an agent-based simulation tool), applied to some specific scenarios.

We have found that there is very limited existing literature that addresses the two topics highlighted above, perhaps in the case of IEDs because the IED problem is a relatively new phenomenon, and in the case of co-evolution methods because the state of the art is only beginning to reach that level.

Therefore, we have taken the approach of providing an overview of the existing CRT work, emphasising those parts that, in our opinion, **may best be applied to IEDs, possibly using co-evolution methods, in the future**. To assist our exposition, in the next section, we describe an example scenario concerned with IEDs, and refer to this scenario as needed.

We then provide a background section, introducing Computational Red Teaming, and some of the technologies that it uses. We then review existing literature on CRT, and provide a summary of commonly used computational tools that the work described in the literature has used.

We conclude with a section that summarises the current situation, and suggests promising future directions.

IMPROVISED EXPLOSIVE DEVICES (IED’S) – AN EXAMPLE SCENARIO

In order to make our discussions more concrete, we would like to introduce an example Improvised Explosive Devices (or IED’s) scenario. Our (fictitious) scenario focuses on the initial When/Where/How layer. First, some definitions are in order:

An IED is “(DOD,NATO) *A device placed or fabricated in an improvised manner incorporating destructive, lethal, noxious, pyrotechnic, or incendiary chemicals and designed to destroy, incapacitate, harass, or distract. It may incorporate military stores, but is normally devised from nonmilitary components. Also called IED*” [1]. Furthermore, according to [2]:

“An IED can be almost anything with any type of material and initiator. It is a “homemade” device that is designed to cause death or injury by using explosives alone or in combination with toxic chemicals, biological toxins, or radiological material. IEDs can be produced in varying sizes, functioning methods, containers, and delivery methods. IEDs can utilize commercial or military explosives, homemade explosives, or military ordnance and ordnance components.

They are unique in nature because the IED builder has had to improvise with the materials at hand. Designed to defeat a specific target or type of target, they generally become more difficult to detect and protect against as they become more sophisticated."

IEDs are an increasing and significant problem. In, [3] Brigadier P.D.Winter, CSC, Commander ADF Counter-IED Task Force concludes:

"In an age of increasing asymmetric warfare, the IED is a weapon of choice for anti-coalition forces, particularly in Iraq and Afghanistan. The enemies which ADF personnel are currently facing overseas are resourceful and adaptable. The situation being faced is not dissimilar to that of the insurgent and guerrilla warfare campaigns of earlier wars. The differences, however, come in the rate of IED attacks and new technologies."

We propose to use, as an example, a scenario in which a defensive force (the blue force) is charged with protecting against and responding to a possible IED attack on a marketplace. Suppose that in this scenario, it is known that a local terrorist cell intends to attack the marketplace, and that recent intelligence suggests that two suicide bombers have been recruited to carry out the attack.

Let us place ourselves in the position of the blue force leaders, who must devise a plan to defend the marketplace as well as possible, given limited resources. Suppose you have a squad of 6 soldiers available for the task. Suppose also that the marketplace is quite large – large enough and with enough obstructions, so that it is not possible to guard all entrances to the marketplace, or to visually observe the whole marketplace from fixed positions. See Figure 1 Figure 6.

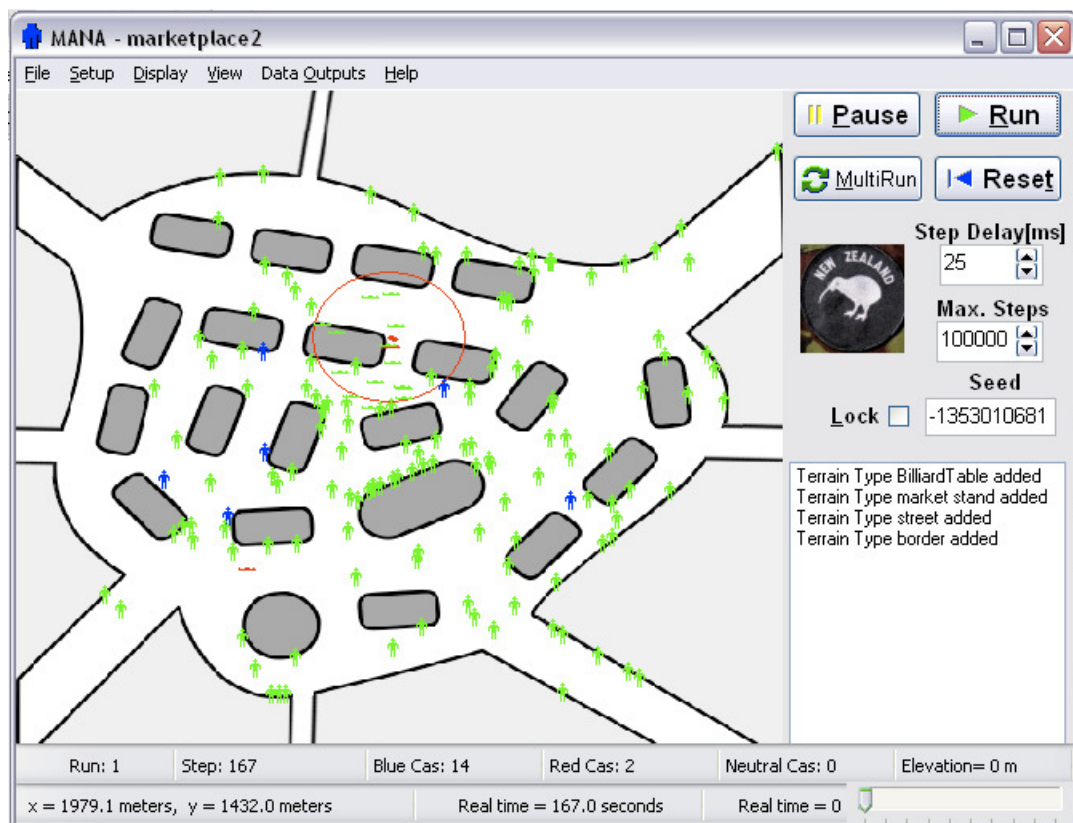


Figure 1 - MANA screenshot of the example IED scenario. There are 8 entrances to the marketplace. Civilians are shown as green icons, there are 6 agents in the blue team, one red agent has been killed (bottom left) and another has just detonated his device (the red circle), killing a number of civilians.

From previous experience, you expect that the bombers will attempt to enter the marketplace via one of the entrances, and mingle with the civilian shoppers, before making their way to a chosen

location at a chosen time and detonating their device. Meanwhile, civilians will also enter the marketplace, visit some of the stalls, and then leave.

Your soldiers could be positioned in strategic locations around the marketplace, or carry out predetermined patrols, or perhaps patrol at random. If a soldier sees someone he or she thinks is suspicious, or perhaps just at random, he or she can challenge that person, ordering them to stop, drop anything they might be carrying, or remove outer layers of clothing etc, and then approach the person and carry out a search. A civilian could be expected to cooperate, while a bomber might immediately detonate their device, or attempt to flee – at which point the soldier is permitted to use deadly force to stop the person.

If one bomber succeeds in detonating their device, you expect the civilians in the area to run away from the explosion and towards one of the exits from the marketplace. One tactic that the bomber might employ is to have the second bomber wait at one of the exits, and detonate at a time chosen for maximum casualties. What should the blue forces do at this time? Should they direct the civilians to a particular exit? Should they try to find the second bomber?

Now the blue force leaders might create a defensive plan using their experience and intuition. How could their plan be tested for weaknesses? Of course, the blue leaders will do their best to anticipate and correct flaws in their plan, but they may easily overlook possible attack strategies that have not occurred to them for some reason – perhaps an unwarranted assumption or blind spot in their thinking. One option would be to create a “red team” consisting of people who did not take part in devising the plan, and charging them with finding an attack plan to defeat the planned defences. But this team too may have its blind spots, and a human team could only consider a few possibilities within a practical timeframe.

This is where a computer simulation can assist. By creating a realistic enough simulation of the scenario, including the physical layout, the likely actions of the human players, the effects of an explosion or gunfire etc, then the potential or likely outcome of a particular red team plan could be assessed. In fact, many simulations could be executed in a reasonable amount of time. A computer system could be created to generate and evaluate many variations on a basic red team plan – for example, varying routes, and the timing and location of detonations. If the blue team plan does not stand up to this analysis, a new blue team plan could be developed that addresses any weaknesses that are discovered. This new blue plan could then undergo further testing and analysis, and so on. One might even imagine a computer system that automatically generates variations of a flawed blue plan, searching for one that is harder for the red team to defeat.

Although it is unlikely that the blue leaders would accept a new blue plan emerging from such a process uncritically – a simulation is just a simulation after all and can never capture every aspect of reality with absolute fidelity – the insights gained from the exercise could provide invaluable guidance.

Indeed, this, in a nutshell, is what computational red teaming is about – using computational models of scenarios in order to gain insights into the strategic strengths, weaknesses and possible alternative courses available to each side in a conflict situation. We will return to this scenario as an illustrative example when discussing recent research and computational tools for computational red teaming in this report.

BACKGROUND ON COMPUTATIONAL RED TEAMING

As mentioned earlier, by *Computational Red Teaming*, we mean any activity involving the use of computer simulations or computational models as a tool in red teaming.

One popular computational red teaming approach is to combine agent-based simulations with computer-based optimisation. In this paradigm, a scenario is analysed through the process of searching for good strategies for one side or the other. This has also been called *Automated Red Teaming (ART)*.

For example, if the blue strategy is predetermined, then a good red strategy is one that is able to exploit some weaknesses in the blue strategy. This might be called the traditional red teaming task – the weaknesses found by the red team can then be addressed to improve the blue strategy. One way to find such a good red strategy is through a search-based optimisation method. A suitable search space is defined (which delineates the possible red strategies), and a search algorithm searches this space for high quality red strategies. The quality of a red strategy is measured by running simulations to answer the question: “What would be the likely result if blue used this strategy and red use this one?”

The process described above could be viewed as a search process itself: one searches for a good blue strategy by positing an initial blue strategy, and finding a red strategy to defeat it. The blue strategy can then be improved to take account of the red strategy. This new blue strategy can then be tested in the same way. Thus the blue strategy can be improved and made more robust, by incremental improvement. It should be noted, though, that things might not play out this way in practice: adjusting a blue strategy to counter a particular red strategy may in itself create new weaknesses that can then be exploited by some other red strategy. There is no guarantee that the “improvement” process will converge. Care must be taken to avoid the pitfalls of this naive approach.

Aside from searching for an optimum as in the ART approach or similar, an alternative is to carry out a carefully designed set of experiments, in which multiple simulations are run using a range of parameter settings for the blue and red team strategies. Statistical methods can then be used to analyse the results to understand parameter effects and interactions. This approach is known as *Data Farming*, a name inspired by the term *Data Mining*. Typically, because of the number of simulations required for statistical testing, some kind of high-performance computing facility, such as a cluster, is needed.

AGENT-BASED MODELLING AND SIMULATION

An agent-based model is a computational model that simulates the actions and interactions of many individually modelled entities (agents). Such models have been used to study so-called “complex systems”.

A complex system is one that is composed of many simple interacting parts, in which overall patterns of behaviour “emerge” by a process of “self-organisation”, where local interactions governed by simple rules result in unanticipated coherent phenomena. Examples of complex systems include financial markets, ecosystems, weather systems, social networks and so on. A well known example is the flocking behaviour of birds. Each bird in a flock follows simple rules that determine its direction and speed of flight, depending on how near or far it is from its neighbours in the flock, and on their direction and speed. The flock as a whole appears to be moving in a coordinated way, but there is no central coordination or control.

In an agent-based model of a complex system, the individual parts are “agents” – individually modelled entities that perceive their environment (using sensors) and act upon it (using actuators).

Usually the individual parts are considered to possess some level of “intelligence”, and Artificial Intelligence techniques are used to model them as “intelligent agents”.

Arguably, a combat situation can be viewed as such a complex system. The interacting parts could be individual soldiers, or squads, or autonomous vehicles and so on. Their interactions are determined by orders and rules of engagement. This is the rationale for using agent-based models or “multi-agent systems” to simulate combat scenarios, as in computational red teaming.

Up to the present time, most computational red teaming work has focussed on tactical combat type scenarios, but with suitable agent-based models, there is no reason that the same analysis paradigms could not be applied to study other situations where a suitable computational model is available. For example, a computational model of cause and effect in a social network such as a terrorist organisation could be used.

It is worthwhile to consider some of the choices available for programming agents for an agent-based system, and how well they fit the CRT context. For this, we follow the exposition of intelligent agents given in Chapter 2 of [67].

Intelligent agents are usually judged according to some performance measure – some measure of how well the agent performs its expected role. A *rational agent* is one that always performs actions that can be expected to maximise this performance measure, given what the agent thinks it knows about its environment. This is a kind of ideal performance that agents typically do not achieve. An example of a feature that a rational agent should have is *autonomy* – the ability to deal with unexpected circumstances, such as being given incorrect or incomplete information about its environment. Achieving autonomy may require the agents to explore its environment, and to have the ability to learn. Rationality can also require an agent to be able to reason perfectly about future consequences of its actions, and to plan for the future. These are high-level competencies that are difficult to achieve in all but the simplest of environments, and generally are expensive in terms of the computing resources required to implement them.

Design choices for intelligent agents depend on properties of the environment that they are meant to interact with. For example, is the environment *fully observable* by the agent (can the agent know everything about the current state of the environment). Clearly, in a combat scenario or an intelligence situation, this is not the case. Is it a *single-* or *multi-agent* environment. Most, if not all, examples of interest for CRT are multi-agent, and this adds complications – interacting with other agents, which may also be rational to some degree or other – cooperating with some (those on the agents own side) and competing with others (those on the other side).

Combat environments are *stochastic* rather than *deterministic* – i.e. chance events affect the outcomes of actions. This also adds complexity. For example, how could a rational agent estimate its future performance, when the outcome depends on chance? In a combat scenario, is it appropriate to consider the average case, or the best case, or the worst case? or something else?

Is the environment *static* or *dynamic* – can the environment change while the agent is thinking about what to do next? Is the environment *discrete* or *continuous* – for combat perhaps the natural choice would be continuous, if we want to model troop movements and speeds etc.

Depending on the nature and complexity of the environment and the task, and depending on the degree of fidelity required of the model, a more or less capable (i.e. more or less rational) agent architecture is called for. In ascending order of rationality, it is usual to classify agents as

- Simple reflex agents – these simply react to the current states of its sensors, possibly using a collection of rules e.g. if shots are fired at me, return fire.
- Model-based reflex agents – these are like simple reflex agents except that the action can be based on the agent's current perception of the environment, which can include previous sensor information and assumptions.
- Goal-based agents – these agents have goals that they try to achieve, and can reason about how to achieve those goals by choosing what to do now and planning what to do next.
- Utility-based agents – these agents try to choose actions that will maximise their performance measure.

A BDI agent (beliefs-desires-intentions) is an example of a type of agent that fits high on the rationality scale [69]. Based on its current beliefs (about the environment and itself, updated with experience), the agent chooses suitable desires (goals), and then forms intentions (plans) to achieve those desires. This is a complex model based on theories of how human reasoning operates. Complex models like this have been used in military simulation systems (e.g. [70]), but not to any great extent in the red teaming context.

AGENT-BASED DISTILLATIONS

Agent-based distillations (ABDs) are low-resolution agent-based simulation models, developed to explore combat scenarios. Agents are reflex agents following simple rules, the environment model is also simplified, often being only 2D and modelled as a grid, and there is no realistic simulation of physics. This allows many simulations to be executed in a relatively short time. Many ABD models have been created: ISAAC, EINSTEIN, MANA, CROCADILE, WISDOM, and Pythagoras to name a few. These and others are described later in this report.

There is a trade-off here: ABDs discard detail in favour of execution speed, and there is a danger that a distillation model may neglect important features of a scenario, resulting in invalid conclusions being drawn. This is the issue of model validation and verification. However, according to [16], for example, increased fidelity does not necessarily increase the accuracy of a model, especially if detail is not consistent with the model's abstractions. The authors note that the creating a distillation-based model at an appropriate level of abstraction, specifying assumptions, and exploring parameters are in themselves beneficial for decision makers.

This issue is discussed further in [46], in which the authors present several distillation studies using MANA and Pythagoras, and use the results in order to design smaller, more focussed experiments using a high fidelity simulator (JANUS). Thus agent-based distillation is to be seen as a part, even if a valuable part, of a larger process.

OPTIMISATION METHODS FOR CRT

The ability to run many simulations of a scenario using agent-based simulation allows us to make use of optimisation algorithms to search for best or worst cases. This can help us to understand the best strategies to use, and also to understand our own weaknesses.

To use optimisation to investigate strategies for a given scenario, we identify a set of strategy parameters that can be varied – this defines the search space. We also define some figure of merit or *measure of effectiveness* or objective that we want to maximise or minimise. The problem is then to find the

parameter values that maximise or minimise this figure of merit. Sometimes more than one objective must be optimised simultaneously, resulting in the need to consider trade-offs.

If a scenario can be described precisely as a collection of equations, it may be possible to apply analytical optimisation methods to obtain exact solutions with relatively few computational demands. However, an agent-based simulation model is not susceptible to this kind of approach. Such models are typically complex systems, in which macro-phenomena arise from many low level interactions and interdependencies, becoming apparent only when the system as a whole is simulated. In addition, as outlined above, typical CRT applications are stochastic, dynamic and not fully observable.

In a case like this, various heuristic search methods can be considered, such as simulated annealing, TABU search, and a number of population-based search methods taking their inspiration from natural processes.

Population-based methods maintain a set of alternative solutions to the problem being considered. This gives them certain advantages when used in decision support. For example, the decision-maker often prefers to be presented with a set of options rather than a single solution, so as to take into account intuition and intangibles not represented in the simulations.

Also, many real-world optimisation problems are multi-objective: there is more than one important factor to be optimised. For example, in a battle scenario, there may be an overall strategic objective to be achieved, and at the same time, it is desirable to minimise friendly casualties. Some optimisation methods require these sometimes-conflicting objectives to be combined into a single overall figure of merit or measure of effectiveness. With population-based methods there are other alternatives.

One alternative is to use the concept of Pareto-optimality: a problem solution is Pareto-optimal if it is not possible to improve it in any one objective without making it worse in some other objective. The problem then does not have a unique best solution, rather the set of all Pareto-optimal solutions forms a trade-off surface (see Figure 2). The aim of a multi-objective optimisation algorithm is to find a representative sample of solutions on the trade-off surface. Population-based methods can do this efficiently by managing a whole set of solutions in parallel.

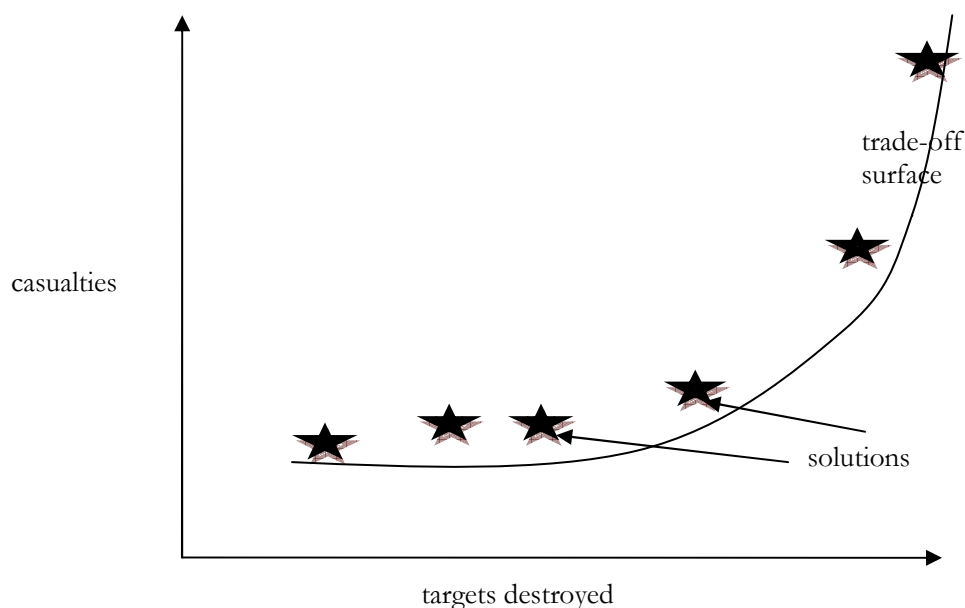


Figure 2 - Trade-off surface for a multi-objective problem. Solutions approximate the trade-off surface. In this example, the aim is to maximise targets destroyed, while minimising casualties. For points on the trade-off surface, it is possible to increase targets destroyed, but only at the expense of more casualties (and vice versa).

EVOLUTIONARY ALGORITHMS

Evolutionary algorithms are population-based global optimisation algorithms based on the principles of natural evolution. In natural evolution, organisms may be viewed as “solutions” whose objective is to perform well in their natural environment. Performing well is equated with producing many well-performing offspring – which usually requires surviving long enough to reach reproductive age, attracting mates, and having sufficient resources to produce and rear new generations of offspring. Individual that perform well in this sense are deemed “fitter”. The offspring of an organism inherit genes from them (slightly modified by mutation and/or, in the case of sexual reproduction by crossover). Since fitter individuals pass on more of their genes to the next generation of organisms than their less fit contemporaries, the proportion of “good” genes in the population increases over time, so that organisms become more fit for their environment.

Evolutionary algorithms mimic this process, by simulating the evolution of a population of organisms (solutions), each being determined by its genes (parameters), and where the number of offspring generated by a organism depends on how well it solves the problem (its “fitness”). Artificial versions of mutation (changing parameter values) and crossover (combining parameter values from several solutions) are designed to ensure an adequate coverage of the parameter search space.

Although there are several main types of evolutionary algorithm (genetic algorithms, evolution strategies, evolutionary programming, genetic programming), and many variations of each of those, the basic principles in each case are similar:

1. Create an initial population of solutions (perhaps randomly)
2. Evaluate each solution
3. If the termination condition is not met
 - a. Select solutions as “parents” based on fitness
 - b. Create a new generation of offspring from the parents, using mutation and crossover
 - c. Evaluate the offspring
 - d. Select the survivors from amongst the parent and offspring
 - e. Continue at step 3

Multi-objective evolutionary algorithms (MOEAs) are mostly based on Pareto-optimality. They follow the same outline as above, except that

- “Fitness” is determined by how “far” from being Pareto-optimal the solution is. There are various schemes for making this determination.
- It is necessary to include a “diversity” mechanism, to ensure a good spread of solutions over the whole extent of the trade-off surface.

CO-EVOLUTION

If one is investigating red team strategies to defeat a known blue team strategy, then it is relatively straightforward to determine the fitness of a candidate strategy – one can simply run the simulation model against the known blue team strategy, and obtain values for the appropriate measures of

effectiveness. If the simulation is stochastic, multiple simulations can be run and average (or some other summary) figures can be calculated.

However, if the blue team strategy is unknown, how can we proceed?

One method is to simulate the red team strategy against a number of different blue team strategies, and hope that a red team strategy that does well against them all will be robust enough to do well against any blue strategy.

Another approach is to mimic a process that occurs in evolution in nature – co-evolution. In co-evolution, two (or more) species evolve together, in such a way that the fitness of an organism in species A depends on how it interacts with members of species B, and vice versa. Co-evolution can be competitive, as in the case of a prey species like a gazelle evolving against a predator species like a cheetah. Or it can be cooperative, as in the case of an insect species evolving against a flowering plant (the insect gets nectar while the plant has its pollen spread, enabling reproduction).

In a red teaming setting, we can simulate competitive co-evolution, by evolving red team strategies against blue team strategies. When we do this, what we hope for is a kind of “arms race” like the gazelle and cheetah (each driving the other to run faster and be more agile), with both red and blue strategies becoming stronger and more robust over the course of the co-evolution.

Unfortunately, this simple and attractive idea has some problems, which must be understood and catered for. When implemented as a kind of evolutionary algorithm in a naive way, a number of “pathologies” can occur, preventing the algorithm from finding high quality solutions. Here is an outline of a naive two population co-evolutionary algorithm

1. Create two initial population of solutions (perhaps randomly), for A and B
2. Evaluate
 - a. each solution in A by testing against the solutions in B
 - b. each solution in B by testing against the solutions in A
3. If the termination condition is not met
 - a. Select solutions from A as “parents” based on fitness
 - b. Create a new generation for A from the parents, using mutation and crossover
 - c. Select solutions from B as “parents” based on fitness
 - d. Create a new generation for B from the parents, using mutation and crossover
 - e. Evaluate the A offspring by testing against B
 - f. Select the survivors for A from amongst the parent and offspring
 - g. Evaluate the B offspring by testing against A
 - h. Select the survivors for B from amongst the parent and offspring
 - i. Continue at step 3

Amongst the problems that can occur with this naive version are *cyclic dynamics*, *loss of fitness gradient*, and *evolutionary forgetting*.

For example, the problem of cyclic dynamics is the following: Suppose we are co-evolving population A against population B. First, population A discovers that strategy A1 does well against this B population, and so population A fills up with A1 strategies. Next, population B discovers that strategy B1 works well against A1, so population B fills up with B1 strategies. Now population A discovers a strategy A2 that beats B1, so population A fills up with A2. Next, population B finds that strategy B2 is good against strategy A2. Finally, population A finds that strategy A1 is good against B2, so A fills up with A1 again. Now a cyclic pattern emerges: A finds A1, B counters with B1, A switches to A2, B goes back to B2, and now A finds A1 again and so on.

This is reminiscent of the child's game, RoShamBo or Rock/Paper/Scissors. In this game, the two players simultaneously choose one of the strategies Rock, Paper or Scissors (usually indicated with hand signs). The winner is decided by the rule Rock beats Scissors, Scissors beats Paper, Paper beats Rock. What is the best strategy for this game?

The answer is complicated because your success depends on your opponent's choice as well as on your own. If you know your opponent's choice, the game becomes easy. For example, if you know he will play Rock, you play Paper. Perhaps the best you can do is to make yourself unpredictable, by choosing randomly. What are the best probabilities to use? Should you choose, say Rock 4 times out of 10, Paper 3 times out of 10, and Scissors 3 times out of 10? No – because if your opponent chooses to play Paper all the time, he will win 4 times out of 10, draw with you 3 times, and lose only 3 times out of 10.

It turns out that the only “safe” way to play is to choose each of Rock, Paper and Scissors $1/3^{\text{rd}}$ of the time. In the long term, no matter what your opponent does, you will win just as often as you lose. If both players settle on this strategy, then this is an example of a “Nash equilibrium” – neither player can get a better outcome by unilaterally changing his strategy. This kind of analysis is the domain of *mathematical game theory*.

In recent years, researchers have applied game theory concepts to better understand the dynamics and pathologies of co-evolutionary algorithms, and have developed improved algorithms that can overcome problems such as cyclic dynamics. The choice of an algorithm for a particular problem depends on the nature of the problem and the kind of “solution concept” that is being sought – for example, whether we are seeking a Nash equilibrium, or a dominant strategy, or a Pareto-optimal strategy.

While co-evolutionary algorithms have been used (with mixed success) for some time to evolve strategies for entertainment games like board games, card games and video games, there has to date been very little work in using co-evolutionary algorithms for Computational Red Teaming. The only work that we know of to date has been carried out as part of the ART (Automated Red Teaming) project – see later.

We believe that there is a great deal of potential for developing co-evolutionary algorithms for use in Computational Red Teaming applications, by applying and improving upon the more modern co-evolutionary algorithms based on game theoretic concepts.

SWARM-BASED OPTIMISERS

Another family of population-based optimisation algorithms inspired by nature is the swarm-optimisers. These optimisers try to imitate the way swarms of relatively unintelligent animals cooperate

to solve difficult optimisation problems. Well known examples are various kinds of Ant Colony Optimisation (ACO) algorithms and Particle Swarm Optimisation (PSO).

ACO algorithms are based on our observation of the way ants cooperate on tasks such as foraging for food. Each ant follows simple rules of movement that take into account previously successful moves made by previous ants. In nature, ants leave trails of chemical (pheromones) when they move, and following ants can detect them, thus providing the means for communication and coordination between ants. ACOs have been applied to a wide range of optimisation problems, but not as yet, as far as we are aware, to CRT tasks.

Particle Swarm Optimisation takes its inspiration from flocks of birds, shoals of fish and swarms of insects, and the way that they move in a coordinated way when, for example, looking for suitable food sources. Individuals coordinate their movements with their neighbours, with the whole swarm achieving a form of emergent coordination.

This is simulated in a PSO algorithm, with individual points (called “particles”) in a problem’s parameter space represented as members of a swarm. Each particle has its own position and velocity, which are updated using simple rules of local interaction, depending on the position and velocity of its neighbours.

Experience shows that PSO can perform better than evolutionary algorithms on some problems (and conversely on other problems), and PSO has similar advantages to evolutionary algorithms, so can be a useful additional tool to have available. The ART project (see below) has experimented with PSO in a CRT context.

DATA FARMING

The term *Data Farming* refers to the CRT method in which “experiments” are designed to gather data from a large number of agent-based simulations, which are then analysed using various statistical and machine learning techniques, to obtain useful knowledge and insights about factors affecting the scenario being studied. The name is intended to be analogous to *Data Mining*, which is somewhat similar, except that the data to be analysed is thought of as already existing, like a natural resource, whereas with data farming, we can specify what data is collected, and then “harvest” the results.

Some of the analysis techniques available include regression, surface fitting, regression trees, standard visualisation methods such as box and whisker plots, bar charts, surface plots, contour plots etc, as well as any other techniques traditionally used in data mining, such as neural networks, decision trees, cluster analysis and so on. When many levels of many parameters are to be explored, efficient experimental designs are needed. An almost ubiquitous kind of design used in the data farming community is Nearly Orthogonal Latin Hypercubes (NOLH). A very readable discussion of the data farming process can be found in [18]. The use of NOLH for designing experiments to analyse high-dimensional simulation models for defence applications was introduced in [19]. Spreadsheets containing orthogonal and nearly orthogonal designs can be found at <http://harvest.nps.edu/software.html>.

LITERATURE SURVEY ON COMPUTATIONAL RED TEAMING

In this section, we review publications describing recent research in computational red teaming. Since it would not be practical to review all recent events, we provide an overview along with more detail on some papers, and direct the reader towards several publication sources with specific output in CRT.

A significant amount of computational red teaming research has been presented in recent years at the Winter Simulation Conference (WSC). The WSC is a major international conference for disseminating recent advances in discrete-event and combined discrete- and continuous-simulation. It has run annually since 1995. Full papers since 1997 are available and searchable on the INFORMS Simulation Society website (<http://www.informs-sim.org/>). Some relevant papers are highlighted below: [3, 9,10,11].

Another regular meeting is the International Data Farming Workshop series, which runs approximately bi-annually. The Data Farming Workshops arose following the completion of Project Albert, a project to develop the data farming concept for the US Marine Corp's Warfighting Laboratory [13]. A history of the project is described in [10]. These workshops are run by the SEED Centre for Data Farming, which is run by researchers based at the Naval Postgraduate School in Monterey, California. The SEED Centre's mission is to:

Advance the collaborative development and use of simulation experiments and efficient designs to provide decision makers with timely insights on complex systems and operations.[12]

The acronym SEED (Simulation Experiments & Efficient Designs) reflects the fact that the main interest of the organisers is in gaining insights on simulations results using efficient statistical methods, but over recent workshops, there have also been regular sessions devoted to the ART framework, using advanced optimisation algorithms combined with simulation. Information on previous Workshops as well as Scythe, the proceedings of the workshops, starting from Workshop 13 in November 2006, are available at [12]. These proceedings document the outcomes of team-based projects carried out at the workshops to examine problems of particular significance at that time. A number of the papers reviewed here are not peer-reviewed academic papers, but team reports published in Scythe [6,8,14,25,26,50,51,52,53,54].

Aside from the work reported in Scythe, a number of research groups have been active in developing their own tools and applying them to red teaming. WISDOM (Warfare Intelligent System for Dynamic Optimization of Missions), for example, was described and used in [28] to investigate the characteristics of the fitness landscape in a combat scenario. They argue that an understanding of the fitness landscape can be used to choose a suitable search algorithm for red teaming applications. WISDOM and its successor WISDOM-II were developed at the Artificial Life and Adaptive Robotics Laboratory at the Australian Defence Force Academy.

In [68, 63], the researchers introduce their Network-Centric Multi-Agent Architecture (NCMAA), which has a number of advantages for modelling complex systems – validation using network concepts; reasoning at the group rather than agent level, and easier parallelisation. This architecture was used as the basis for WISDOM II. Combat scenarios can be developed using networks to define command and control structures, visibility of agents, communications, situational awareness (maintaining each agent's view of the world) and a firing network. Actions available to agents depend on their roles, and work through these networks. In the papers, WISDOM-II is demonstrated using a scenario in which modelling of communications is critical. In [34], WISDOM-II is used in a red teaming study for risk assessment. Using a ratio of red casualties to blue casualties as the figure of merit, they ran simulations comparing the performance of a traditional blue force with one using network-centric warfare, and investigated the effects of parameters such as force size and communication range.

In [65], the authors used NSGA-II, a commonly chosen multi-objective evolutionary algorithm, to investigate a land combat scenario simulated using WISDOM-II. In their experiments, they evolved the blue force parameters to minimise cost and casualties. The main purpose of the study was to investigate

the effect of parameter setting for the evolutionary algorithm. They found that the optimisation results were not very sensitive to these parameters.

A recently described agent-based simulation model for red teaming is FAMES (Fully Agent-based Modelling and Emergent Simulation) from Korea [61]. This model uses an evolutionary process to search for suitable tactics in a given scenario (the reported scenario is one on one battleship combat). The details of how this evolutionary process works are not clear to us from this paper.

The ART (Automated Red Teaming) tool and the associated ACE (Automated Co-evolution, an extension featuring co-evolution) has been developed by DSO Laboratories, Singapore, in conjunction with collaborators at Nanyang Technological University [47,48,49,44]. The framework has been used in a number of IDFW's [50,51,52].

In [47], ART was introduced and an example Urban Operations scenario was examined. A coastal surveillance scenario and an anchorage protection scenario were studied in [48]. Automatically generated strategies were compared with strategies determined by manual red teaming. ART was found to produce novel solutions not found by the manual teams. The same anchorage protection scenario is studied again in [49], where the idea of “evolvable simulation” is introduced. This is where the search space is modified as part of the evolutionary process (in this case, the number of waypoints is varied.)

The ART architecture consists of a central control module, which can be interfaced with different simulation models (usually MANA), and optimisers (usually an evolutionary algorithm), and can utilise a cluster for carrying out simulations to evaluated strategies, providing fitness values for the optimiser.

In [44], the authors report on experiments with the anchorage protection scenario, comparing the performance of variants of a competitive co-evolution algorithm that is now incorporated in ACE. The algorithm is a simple one, in which the blue force strategies is evolved against a fixed population of red strategies, and then the red strategies are evolved against the (previously evolved) blue strategies, and so on. Several evolutionary algorithms (EPGA – Elite Pareto Genetic Algorithm, developed in-house at DSO, and PSO – Particle Swarm Optimizer), and several solution concepts (best vs best, best vs all) were compared.

The ART methodology is applied to a strategy game scenario in [45]. EVOSIM, was used for this study. EVOSIM is a new implementation of the ART concept, designed for flexibility and implemented in Ruby. In this study, MANA was combined with an MOEA, NSGA-II.

Another potentially interesting source of publications related to CRT is JASSS, the Journal of Artificial Societies and Social Simulation, produced at The University of Surrey in the UK – available online at <http://jasss.soc.surrey.ac.uk/JASSS.html>. An example of a study reported here that might be of relevance to IED's in an indirect way is [72], in which the author uses a NetLogo simulation to model interactions between UN peacekeepers, civilians and insurgents, studying the effect of the number of peacekeepers and the tendance of civilians to scan for insurgents on outcomes. As another recent example, in [73], the authors use a simulation model based on logical rules to describe consequences of actions. The paper studies aspects of the interplay between reputation and “hot-spots” of criminal activity, but the similarities with insurgent activity are obvious. Studies such as these illustrate the potential for agent-based simulation of social networks (the term Agent Based Social Simulation is used (see [74]).

IEDS

There has been relatively little reported computational red teaming research that is specific to IEDs. In this section we summarise several papers reporting initial efforts to apply CRT methods to study IED problems.

In [15], the authors focus on predicting likely future events based on multiple executions agent-based simulations. They developed a tool, DEFUSE, which contains a leadership model (for factional leadership and strategic thinking), a process model (for resources and tasks needed to place an IED), and an environment model (terrain plus overlays for specific properties). “Polyagents” representing teams or other entities generate “ghost” agents that project the teams’ activities into the future. Although this is not a red teaming tool as such, analysts can use it to examine the effects of different courses of action.

In [16], the authors used MANA in conjunction with data farming experiments to examine the use of communications jammers to counter IED threats to ground-based convoys. They examined the effects of jamming effectiveness, length of suppression time, and convoy speed. An unexpected finding was that there was no difference in the effectiveness of ground and air-based jammers in the scenarios considered. The authors commented that although MANA was a useful tool in this context, follow up work should use higher resolution, physics-based simulations, utilising accurate (classified) technical data.

One of the most recent data farming workshops included a task involving the use of agent-based models to study IED scenarios [14]. This particular team focussed on the “Attack the Network” aspect of JIEDDO’s Counter-IED approach (the others are “Defeat the Device” and “Train the Force”). The team used a Pythagoras-based model in conjunction with network analysis and visualisation tools. They list data farming analysis and optimisation using tools such as ART amongst issues identified for future work.

COMPUTATIONAL TOOLS FOR RED TEAMING

In this section, we list and describe some of the simulators and optimisers that have been or can be applied for CRT. There are many agent-based simulation tools now in existence. Here we summarise the features of those that have been commonly applied to CRT tasks.

ISAAC AND EINSTEIN

ISAAC (Irreducible Semi-Autonomous Adaptive Combat) and EINSTEIN (Enhanced ISAAC Neural Simulation Toolkit) were developed by Ilachinski [35,36,37] of the US Marine Corps Combat Development Command, and many subsequent agent-based distillation systems build on this design. The simulation takes place on a 2D map, populated with agents. The dynamics of the simulation are based on mobile cellular automaton rules. Agent actions are determined by

- Doctrine – a default rule set
- Mission – mission-specific goals
- Situational Awareness – an internal map from sensor information
- Adaptability – internal mechanism to change behaviour and rules

Each agent also has a “personality”, defined as a set of propensities to move towards or away from other entities, which are affected by the local surroundings of the agent (distance to objective, number of live enemies in sensor range etc). ISAAC has a genetic algorithm mode, which allows for evolving agent personalities for one team.

ISAAC was later redeveloped and extended as EINSTEIn, which added a graphical interface and visualisation, neural net and machine learning capabilities, and analysis tools.

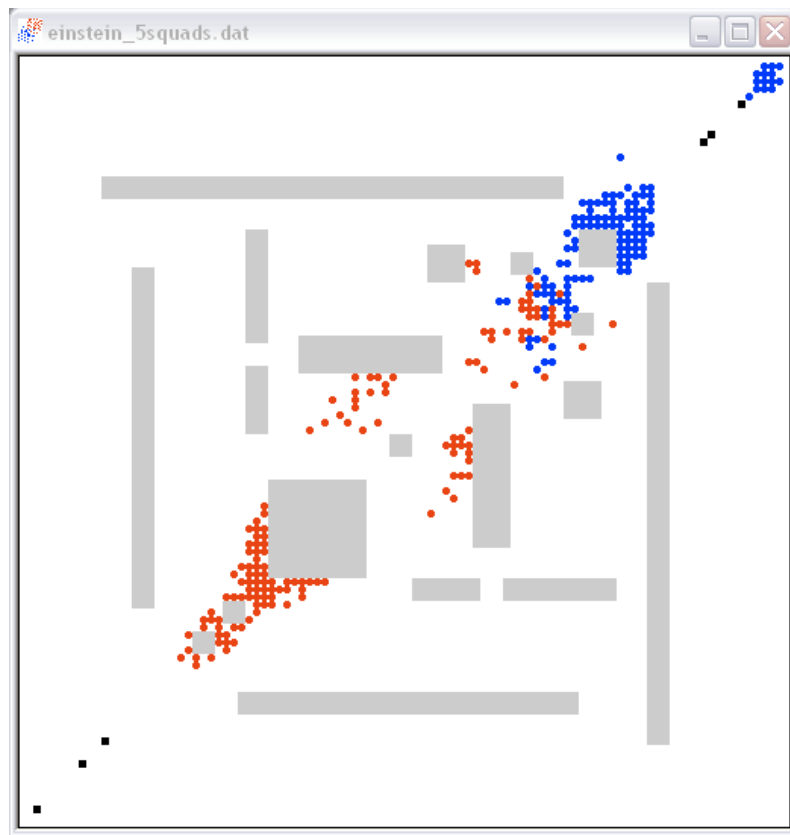


Figure 3 - An EINSTEIn screenshot

EINSTEIn is freely available from [38]. It is Windows-specific.

MANA

MANA (Map Aware Non-uniform Automata) was developed by the Operations Analysis group at Defence Technology Agency, New Zealand [39,40,41,42]. It has gone through a number of revisions, each revision adding improved or additional functionality. MANA is not freely available, and is Windows-specific.

Like ISAAC, it is based on agents situated on a grid-based map. Sensor and weapon characteristics can be specified, and agents are grouped into squads. MANA uses a waypoint system to allow for complex movement patterns to be used. There is support for communications links and information sharing between squads, for modelling network-centric warfare. There is integrated data farming support. There are specific modules for aircraft movement and search. Some simple analysis and visualisation is included, as well as a genetic algorithm module.

Canada's DRDC Valcartier Operations Research Team recently (2006) carried out an evaluation of MANA in order to test its suitability for several simulation studies [43]. They found that MANA has many limitations, in their case making it impossible to use MANA for their studies. For example "careful formation fighting and complex non-scripted interactions among agents" were found to be unobtainable. They proposed a list of 26 improvements that would address many of their issues. They also found workarounds for a number of MANA's shortcomings.

For example, to simulate a coordinated retreat, two fictitious invisible agents were added and equipped with “weapons” that did no damage, but allowed certain personality changes to be triggered. We found that we had to use similar (though not so convoluted) workarounds to achieve some of the behaviours we wanted in the example IED scenario.

To test MANA’s adaptability for an Improvised Explosive Device scenario, we modelled the sample scenario described earlier. MANA seems a priori a reasonable choice, based on the combination of an easy setup and the frequent usage in applications of Automated Red Teaming. Recall the scenario depicted in Figure 1. To prevent the possibility of all entrances being protected by one agent, eight roads to the marketplace were designed, which should be protected by six blue agents. The red team consists of two agents attacking from opposite directions.

A number of workarounds and compromises had to be made to implement this scenario in MANA. For example, the neutral agents were defined to be blue agents without any weapons or reactions toward red agents (as otherwise the red agents would not attack them). To provide a general orientation of the red agents, a waypoint was created in the middle of the marketplace. We wanted the red agents to detonate their devices when there were many neutral agents nearby. As there is no way for an agent to know how many opponents are standing nearby, an alternative mechanism had to be developed. The distribution of the neutrals was arranged to be more crowded in the centre by defining a waypoint, which they slightly follow. The longer distance to the crowded areas is comparable with the time a suicide bomber has to choose a place for denotation. If a blue agent detects a red agent, he will change to a trigger state and try to follow him to eliminate the danger, while the red agent runs away towards the next neutral to explode his device. If no detection occurs, the red agent will denotate after a fixed time in a trigger state, which is activated after seeing the first neutral. The explosion is realised as a short distance gun with a certain blast range, which makes the presence of another agent nearby necessary.

The need for workarounds is not so much a specific criticism of MANA but rather an example of the trade-off between simplicity and fidelity. Similar limitations could undoubtedly be found with most if not all the other distillation-type simulation tools available.

One of MANA’s inbuilt features is a genetic algorithm that may be used to optimise parameter settings for a simulation. The user interface allows the user to select certain parameters to be optimized – see Figure 4. The user can also define the Measure of Effectiveness (MOE) to be optimized – see Figure 5 (the MOE here for the red team is the number of civilian casualties, which rises from an initial value around 13, to around 15-16).

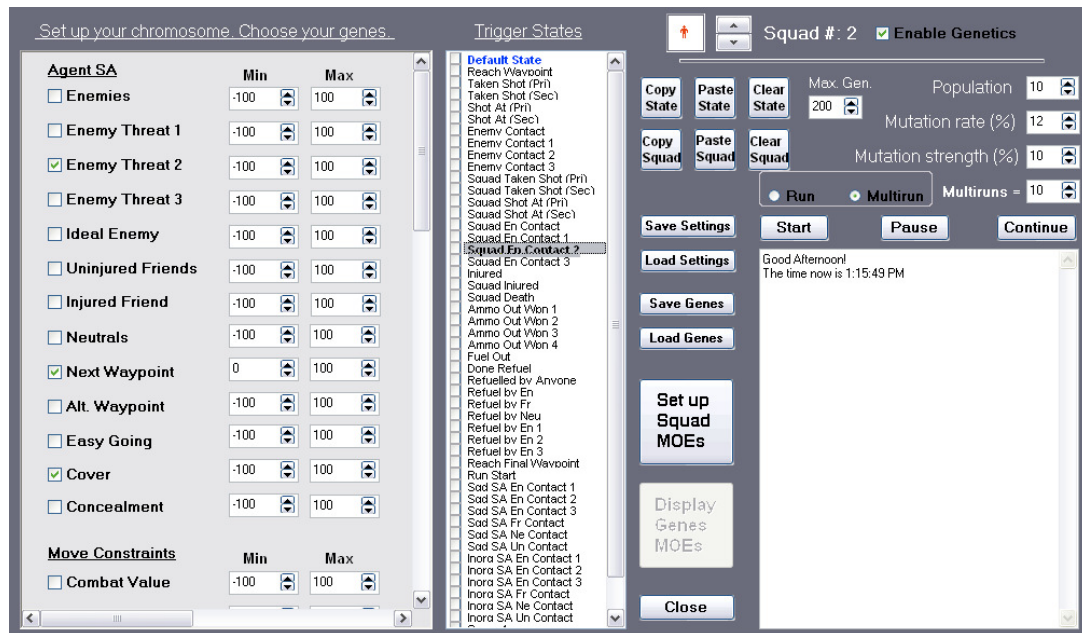


Figure 4 - MANA's user interface, showing the screen that allows the user to select parameters to be optimised by a genetic algorithm.

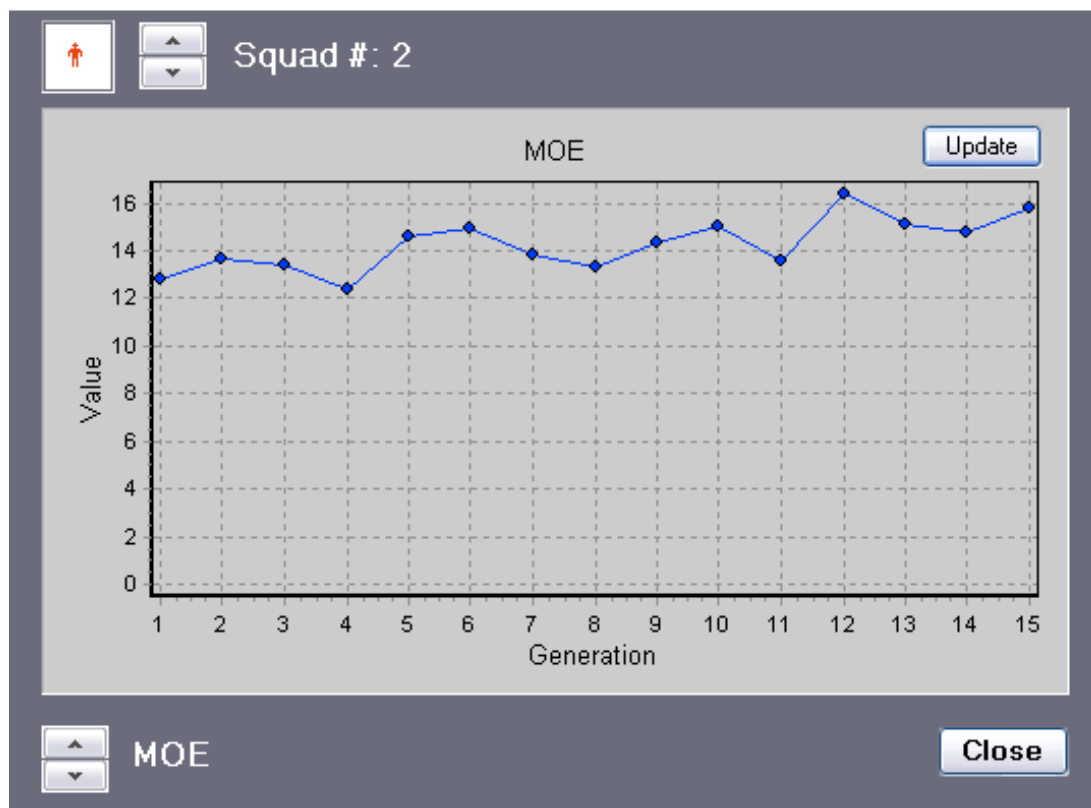


Figure 5 - MANA screen showing improvement in MOE over a number of generations of a genetic algorithm.

PYTHAGORAS

The multi-agent simulation tool Pythagoras was developed to support project Albert, which was implemented by an U.S. Marine Corps-sponsored international initiative [55]. The intention of project Albert was the creation of capabilities for decision makers to analyse a scenario in context with data

farming. Therefore many data points are captured to create landscapes with outcomes of the simulations [56].

Pythagoras enables the definition of personalities for the agents in those scenarios. The concept of soft rules for defining a range for a specific behaviour allows the user to control the agents in a fuzzy way without restricting their variability by setting it to a fixed value. This mechanism also reduces the user input without creating completely identical agents.

The agent's personal desires lead them to move or to shoot at an opponent. The movement can be calculated by a weighted vector algebra between all desires or the strongest two, by priority or randomly with considering weights. Their behaviour can be independent of other agents or supervised by squad leaders. Trigger states can be used for changing the parameters during one simulation under some predefined circumstances. (This is similar to EINSTEIN and MANA.) The definition of a behaviour tree enables the creation of dependencies between events.

The introduction of dynamic affiliation allows the agent to change sides as a result of a specific event or the influence of other agents. This influence among each other is realised by the use of non-lethal weapons. The user can create a communication network, which can include a message for an attribute change additionally to information about other agents.

The fact Pythagoras is written in Java allows usage on many operating systems. In combination with the ability to be run inside a batch job, it is applicable for running on several clusters systems producing more output in the same time, which allows testing a larger number of parameter values or increasing the simulation speed. The current version 2.1.1 of Pythagoras is only available for the US Government and their affiliates and Naval Postgraduate School (NPS) students or faculties.

CROCODILE

CROCADILE (Comprehensive, Research Oriented, Combat Agent Distillation Implemented in the Littoral Environment) was developed at the Australian Defence Force Academy (ADFA) at the University of New South Wales. CROCADILE has a number of distinguishing features [31]:

- 3D or 2D environment in which agents interact.
- Physics model that includes velocity, acceleration, line-of-sight, and collisions.
- Projectile-based physics combat resolution as well as probabilistic resolution.
- Movement by Land, Air, and Water.
- User written agents - Users can code their own agents using any paradigm (e.g., BDI) and load them into CROCADILE.
- Strong OO design - separation of agents from simulation, and separation of agent capabilities from agent behaviour.
- Sophisticated command, mission, and communication structures (networks) for agents.
- Higher fidelity combat resolution models that incorporate blast effects, round penetration, rate of fire, and line-of-sight.
- Database of world objects - scenarios, agents, agent groups, weapons, communication, sensors etc. - that can be saved individually and used for subsequent scenarios.
- Comprehensive simulation event logging. Includes time-stamping of events. Analysis possible by visualiser provided as part of CROCADILE download, or through commercial spreadsheet.
- Multi-team structure including neutrals, levels of alliance/enmity, and communication between teams.

The tool is freely available from the CROCADILE home page [32]. Also available from the same site is TDSS, a Courses of Action analysis tool that works with CROCADILE. CROCADILE is written

in Java (and so is platform independent), and can be extended by adding developed Java classes. Figure 6 shows an example CROCADILE screenshot.

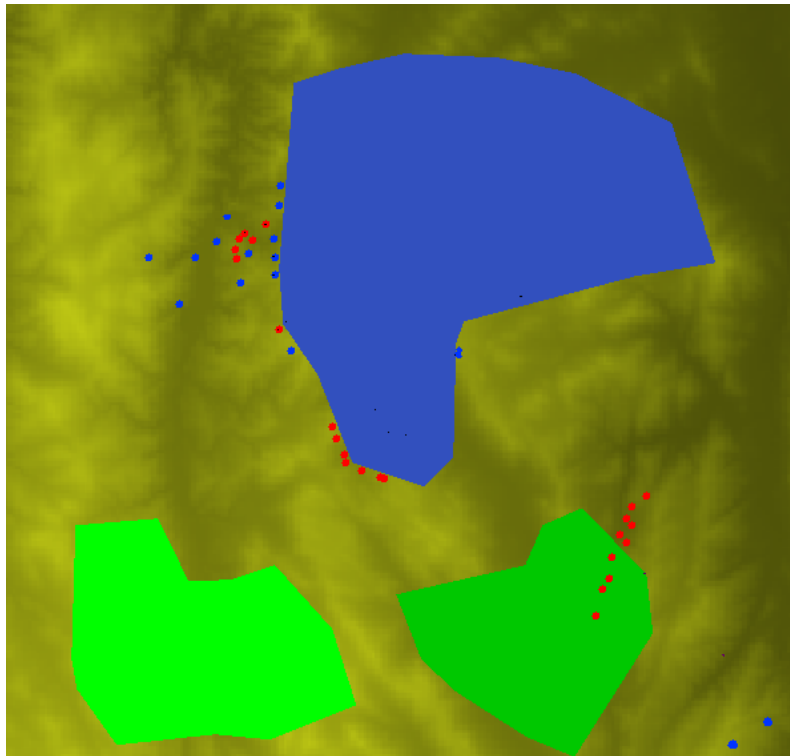


Figure 6 - Screenshot of a CROCADILE scenario. The background is a 3D terrain model, the blue area is a lake, the lower left green areas is a different vegetation type (with modified scanning, movement and firing), and the lower right green area is a different terrain type (with modified probability of sustaining damage, and penetration). The red and blue dots represent red and blue forces.

CROCADILE does not seem to have had much use since it was first developed in 2002, and crashed several times when we tested it, so we are unsure whether it is currently supported.

WISDOM AND WISDOM-II

As mentioned earlier WISDOM (Warfare Intelligent System for Dynamic Optimisation of Missions) [28] and WISDOM-II are multi-agent combat simulation systems developed at the Australian Defence Force Academy. WISDOM agents have sensors, capabilities, movement and communications. Agent movement is determined to “personality” weights, which represent the tendency of an agent to move towards injured or uninjured teammates, opponents and objectives (similar to EINSTEIN, MANA etc).

WISDOM-II is a redesigned and redeveloped “network-centric” simulation platform. It defines five network types, command and control, vision, communication, information fusion, and engagement. Four agent types are supported and each agent has characteristics (health, skill etc) and personality (attraction/repulsion towards other agents). Swarm leaders can make plans and give orders.

WISDOM-II has a number of capabilities distinguishing it from earlier, simple ABD models which go some way to overcoming the limitations of those (too?) simple models.

PAX (AND PAX3D)

The agent based simulation model PAX was developed by EADS Dornier. The development was initiated and sponsored by the German Bundeswehr Training and Doctrine Command (TRADOC) and

assisted by the Operations Research Division of the Bundeswehr Center for Analysis and Studies. It focuses on the issue of Command and Control (2C) [58]. It contains basic concepts for dealing with intra-civilian simulation and the interaction between military and non-military side. The modelling for civilians inside PAX is optimised for Peace Support Operations like analysing de-escalating problems. The soldiers are configured in small groups and follow strictly to certain rules without considering any psychological aspects. This simplifies the analysis of strategies, but doesn't consider any potential side-effects. A representation as a flow chart can be used for designing those rules. PAX concentrates on modelling asymmetric multi-party scenarios [59]. Each individual contains of a number of parameters, which influence their behaviour. The successor PAX3D extended the simulations to a 3 dimensional environment and revised the tactical behaviour of the soldiers [26].

GENERAL PURPOSE MULTI-AGENT SIMULATION TOOLS

Although not specifically designed for Computational Red Teaming, there are a number of multi-agent simulation tools that could be readily applied, although high-level programming skills are required. Some of these are MASON, RePast, Ascape, StarLogo and NetLogo. Here we describe NetLogo and MASON, two that have been used in recent Data Farming Workshops.

The modelling environment NetLogo is designed for modelling natural and social phenomena [57]. The simulation environment is developed by Center for Connected Learning and Computer-Based Modelling. The download, which is available for free, includes many sample simulations. These examples cover various applications in conjunction with natural a social science, like biology, medicine, economics, social psychology, computer science and further more.

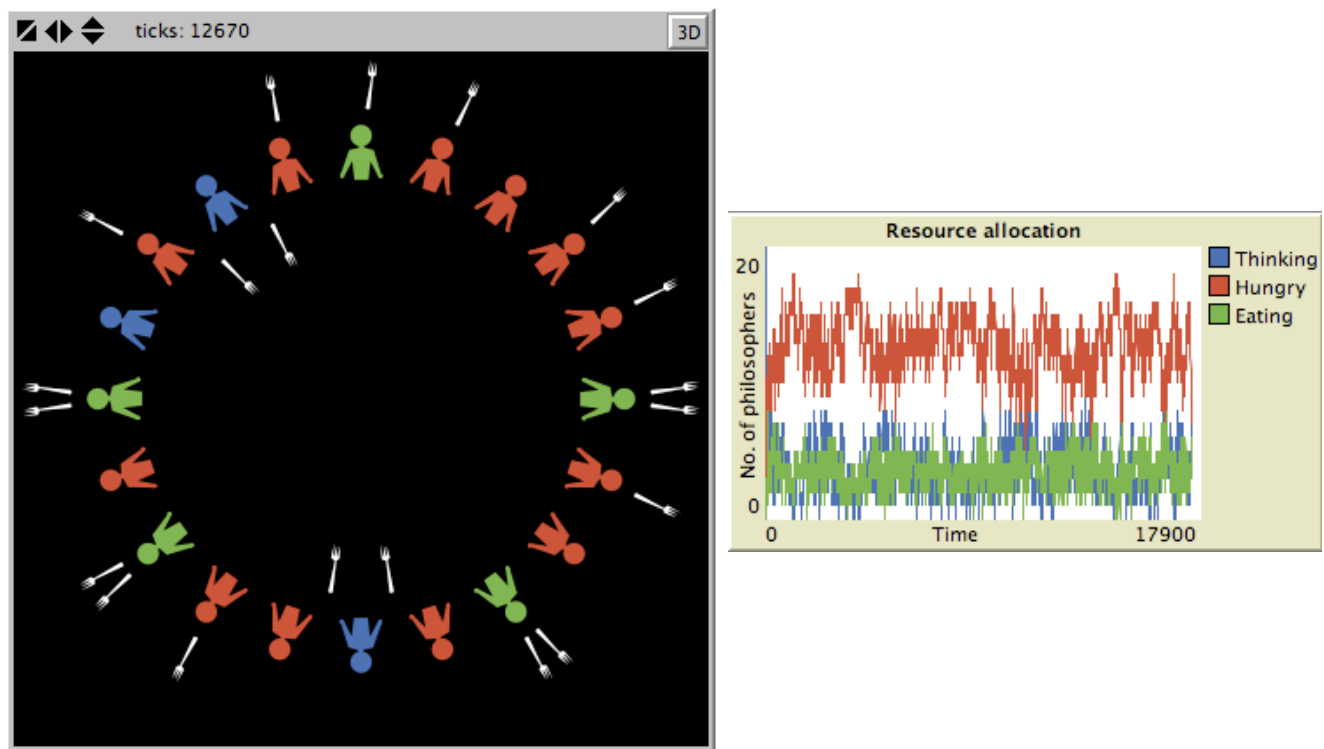


Figure 7 - NetLogo simulation based on the Dining Philosophers problem.

A simulation consists of turtles and a grid of patches. The usage of those two types is completely dependent on the underlying problem. In an IED scenario a turtle for example might represent an agent and the patches the structure of the map. The turtles and patches are programmable with an integrated programming language similar to Lisp. They can communicate with each other and execute tasks

concurrently. Multiple "breeds" of turtles with different parameters can be created for diversification. In contrast to other multi agent simulation tools, the parameters are not limited to a fixed set, as they are interpreted in the programming language. Additionally in each simulation one observer exists as superior instance.

NetLogo is executable on various operating systems, as it is written in Java. An application programmer's interface was introduced to extend a possibility for integrating Java code in the simulations.

Another multi-agent simulation tool, MASON (Multi-Agent Simulator of Neighborhoods), was developed at George Mason University. It is freely available and, like NetLogo, written in Java. It is rather flexible, and includes 2D and 3D visualisation and the ability to produce movies, charts and graphs etc. Adding to its suitability for CRT is ECJ (Evolutionary Computation in Java), an evolutionary algorithm toolkit, written in Java and designed to work with MASON. The toolkit includes various evolutionary algorithms (genetic algorithm, evolution strategies and genetic programming), and includes co-evolution and multi-objective evolution, as well as PSO and some others. It has support for several types of distributed evolutionary algorithms (Island model and Master-Slave). ECJ is downloadable from <http://cs.gmu.edu/~eclab/projects/ecj/> and MASON from <http://cs.gmu.edu/~eclab/projects/mason/>.

Repast is another open-source and widely used agent-based simulation and modelling toolkit originating at Argonne National Laboratories, which has more recently been enhanced as Repast Symphony or Repast S [71]. This is a powerful and extensible system with an architecture that readily supports network-based models suited for modelling physical networks as well as social networks, and useful visualisation tools.

PROTOTYPE OR SPECIAL PURPOSE TOOLS

From time to time a new prototype simulation tools may be developed for a specific type of application. Here are two examples from recent Data Framing Workshops.

JTEAM (Joint Test and Evaluation Agent Model) is a prototype agent-based simulation being developed as part of the JMEDF (Joint Mission Effectiveness support using Data Farming) project, supporting the Netcentric Systems Test Program. It is written in Java using the MASON framework. JTEAM was introduced in IDFW18 where it was used to study a convoy protection scenario and also a casualty evacuation scenario.

LBC is an in-house model developed by TRAC-Monteray, specifically to study logistical supply problems. It is a low-resolution, object oriented, stochastic, and discrete event model programmed in Java and incorporates Simkit. Functionality includes planning and decision support features to enable a simulated sustainment decision maker to monitor the LCOP, forecast demand for most classes of supply, and initiate and adjust missions to distribute supplies and perform sustainment functions. LBC was introduced in IDFW18 [54].

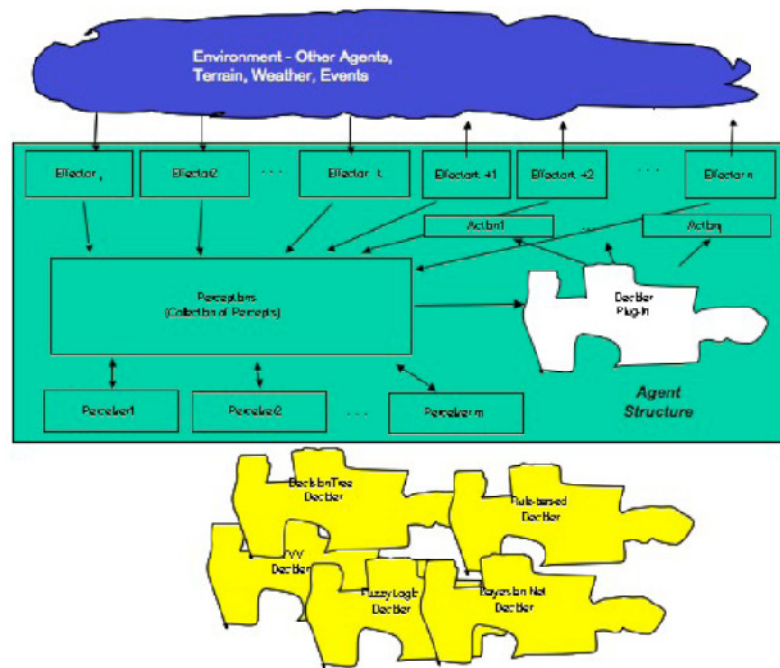


Figure 8 - JTeam architecture, from [53]

A COMPARISON

In [33], the author carried out a review of agent-based models with a view to evaluating them for use in battlefield simulations. Models were evaluated based on a number of general criteria such as flexibility, speed, and documentation, as well as criteria specific to modelling complex adaptive systems such as degree of adaptation (agent capabilities) and self-organisation (e.g. feedback between agents). The models evaluated were BactoWars, EINSTein, MANA, MASON, NetLogo, Repast, Swarm and WISDOM-II.

The three models recommended were MASON for overall capability in modelling complex adaptive systems (with the proviso that high level programming skills are needed), followed by NetLogo (easier for programming). As far as battlefield simulation models go, the recommended system was EINSTein.

TOOLS USED IN RECENT DATA FARMING WORKSHOPS

To provide one view of the most common tools and analysis techniques in use, we surveyed the reports of recent IDFW's and we present here the results in a tabulated form. The most common analysis method is data farming, usually using NOLH designs, but ART (or ACE) were also used, and it is interesting to note an increasing trend in recent IDFW's for the teams to report their intention to supplement data farming with analysis using optimisation and red teaming in the future.

Amongst the agent-based simulation tools, MANA, Pythagoras and PAX appear to be the most consistently used.

Workshop	Simulation tools	Analysis
Scythe 7 – IDFW 19	MANA Arena and Excel PAX3D ABSEM ITSim	Data Farming

	Pythagoras	
Scythe 6 – IDFW 18	MANA (2) JTEAM and Pythagoras TLCM-AT and EXTENDSIM ABSEM PAX3D JTEAM ARENA,NetLogo,Pythagoras,Sandis ITSim LBC PSOM CG	ACE BOLT DOE Data Farming
Scythe 5 – IDFW 17	- not available	
Scythe 4 - IDFW 16	MANA(3) PAX ABSEM LBC TLCM-AT Pythagoras Several in-house agent-based models JDAFS	Data Farming ART PAX Analysis Toolbox
Scythe 3 – IDFW 15	MANA(3) Pythagoras Analytic models PAX POW-ER MASON NetLogo(2)	ART Data Farming Cluster and outlier analysis

CONCLUSION AND FUTURE DEVELOPMENTS

In this report, we have provided background summarised recent research in computational red teaming using agent-based simulations. Currently, most existing tools are distillations, nearly all being based on the ISAAC/EINStein “mobile cellular automaton model”, with a grid representation of geography, and agents whose movement is determined by “flocking” rules. There are two main analysis approaches – data farming as a means to understand the generalised effects of various parameters in a scenario, and optimisation using an evolutionary algorithm or similar heuristic global search algorithm. These methods have been quite successful in efforts to model and analyse a great variety of scenarios, investigating land, sea and air combat, UAV’s, communications, public unrest to name a few.

To date, little has been done specifically with IEDs, but we expect these same techniques to be extremely valuable as a tool to study the various aspects of the IED problem. However, there are some improvements that could be made.

As a number of authors have noted, agent-based distillations trade simulation fidelity for execution speed, allowing thousands of simulations to be run in a reasonable time. As high performance

computing becomes more commonplace and platforms become more powerful, we think that the balance has shifted and it may be appropriate to develop new tools that are better able to model important aspects of real scenarios. A related issue is the inflexibility of simple distillation tools, which sometimes forces the modeller to adopt tortuous workarounds to achieve something resembling a desired effect.

In the area of analysis using optimisation, co-evolutionary methods seem to be an attractive alternative to simple optimisation methods, providing the ability for both sides to “think out of the box”, rather than, for example, the read team making (possibly incorrect) assumptions about the blue strategy.

In the IED application, where the enemy is constantly trying to catch us on the wrong foot, we would argue that greater sophistication and fidelity in modelling, especially with regard to modelling of decision making, and exploration tools using co-evolution to anticipate novel strategies, are two developments that would enhance the value of computational red teaming.

Lastly, computational red teaming has the potential to be applied to a much wider range of scenario types, with applications for counter-IED amongst others. For example, computational models of social networks such as terrorist organisations could readily be combined with analysis tools such as visualisations, optimisation, and statistical studies.

In summary, we see great potential for CRT as an invaluable tool in counter-IED and other defence applications. The current state of the art can be improved in a number of ways:

- More capable agent-based simulation models, including
 - models with greater fidelity, and
 - applicability to a wider range of scenario types, including network-based models suitable for modeling social and logistical networks.
- Better optimisation tools, with the capability for
 - multi-objective
 - human centred
 - co-evolution.
- A comprehensive computation environment supporting a high quality CRT capability.

REFERENCES

1. improvised explosive device, DOD Dictionary,
http://www.dtic.mil/doctrine/dod_dictionary/data/i/10302.html, accessed 23 March 2010-03-23
2. anon, Improvised Explosive Devices (IEDs)/Booby Traps,
<http://www.globalsecurity.org/military/intro/ied.htm>, accessed 23 March 2010.

3. Winter, P., Meiliunas, A., Bliss, S., Countering the improvised explosive devices threat, The Royal United Services Institute of New South Wales Inc, United Service, Volume 59, Number 3, September 2008
4. Chua, C.L., Sim, W.C., Choo, C.S. and Tay, V., Automated Red Teaming: An Objective-Based Data Farming Approach for Red Teaming, Proceedings of the 2008 Winter Simulation Conference, pp. 1456-1462, 2008.
5. Choo, C.S., Chua, C.L. and Tay, V., Automated Red Teaming: A Proposed Framework for Military Application, In Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, pp. 1936-1942, ACM, 2007.
6. Sim, W.C., Choo, C.S., Ng, E.C. and Martinez-Tiburcio, F., Applying Automated Red Teaming in a Maritime Scenario, In Scythe, Vol 1, No 1, pp. 38-39, 2006.
7. Upton, S.C., and Johnson, S.K., and McDonald, M.J., Breaking Blue: Automated Red Teaming using Evolutionary Algorithms, Proceedings in the Genetic and Evolutionary Computation Conference, 2004.
8. Wong, A., Sim, W.C., Chua, C.L., Lim, Y.L., Chin, S., Teo, C., Lampe, T., Hingston, P. and Abbott, B., Applying Automated Red Teaming in a Maritime Scenario, Scythe, 16, 2007.
9. Low, M., Chandramohan, M., and Choo, C.S., Application of a Multi-objective Bee Colony Optimization Algorithm to Automated Red Teaming, Proceedings of the 2009 Winter Simulation Conference, pp. 1798-1808, 2009.
10. Horne, G. and Schweirz, K-P., Data Farming around the World, Proceedings of the 2008 Winter Simulation Conference, pp. 1443-1447, 2008.
11. McDonald, M. and Upton, S., Investigating the Dynamics of Competition: Coevolving Red and Blue Team Simulation Parameters, Proceedings of the 2005 Winter Simulation Conference, pp. 1008-1012, 2005.
12. SEED Centre for Data Farming, <http://harvest.nps.edu/>
13. Project Albert, <http://www.projectalbert.org/>
14. Byers, K., Lesnowicz, E., Meyer, T., McDonald, M., Upton, S., Silwood, N., Engleback, N., Middleton, D., Malcolm, P., Askman, V., Gunzelman, K., DeLange, J. P., Lowry, B., Utility of Distillation Modeling, for Countering IEDs, Scythe, Issue 6 Workshop 18, 2009
15. Van Dyke Parunak, H., Sauter, J., Crossmann, J., Multi-Layer Simulation for Analyzing IED Threats, 2009 IEEE International Conference on Technologies for Homeland Security, pp. 323-330 , 11-12 May 2009
16. Ayvaz, U., Dere, M., Tiah, M., Using the MANA Agent-Based Simulation Tool to Evaluate and Compare the Effectiveness of Ground-Based and Airborne Communications Jammers in Countering the IED Threat to Ground Convoys, Proceedings of the 2007 Spring Simulation Multiconference, Vol. 2, pp. 113-118, 2007

17. Meyer, T.E., Koehler, M., Barry, P. and Tivnan, B. How Simple is Simple Enough? Military Modeling Case Studies, Proceedings of the Agent 2005 Conference, available at <http://www.agent2005.anl.gov/proc2005.html>, 2005
18. Sanchez, S. and Lucas, T., Exploring the world of agent-based simulations: simple models, complex analyses, Proceedings of the 2002 Winter Simulation Conference, pp. 116-126, 2002.
19. Cioppa, T. Efficient Nearly Orthogonal and Space-Filling Experimental Designs for High-Dimensional Complex Models, Naval Postgraduate School Dissertation. Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.99.7853&rep=rep1&type=pdf>
20. Burtsev, M. S., Korotayev, A., An Evolutionary Agent-Based Model of Pre-State Warfare Patterns: Cross-Cultural Tests, World Cultures 15, pp. 17-38, 2004
21. Decraene, J., Zeng, F., Low, Y. H., Zhou, S., Cai, W., Evolutionary Agent-based Simulations applied to Military Decision Making, Preprint submitted to Simulation Modelling Practice and Theory, 9 October 2009
22. Yang, A., Abbass, H. A., Sarker, R., Land Combat Scenario Planning: A Multiobjective Approach, Proceedings of SEAL'06 LNCS, Haifei, China, 2006
23. Zitzler, E., Thiele, L., Deb, K., Comparison of Multiobjective Evolutionary Algorithms: Empirical Results, Evolutionary Computation Vol. 8, No. 2, pp. 173-195, 2000
24. Xu, Y. L., Low, Y. H., Choo C. S., Enhanced Automated Red Teaming with Evolvable Simulation, GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation, pp. 687-694, 2009
25. Lauren, M., Chong, N. E., McDonald, M., Pickl, S., Sanchez, R., Maritime Force Protection Study using MANA and Automatic Co-Evolution (ACE), Scythe, Issue 6 Workshop 18, 2009
26. Seichter, S. Lt., Lampe, T., Schwarz, G., Sim, M. S., Tan, L. M. E., Johnson, R, PAX3D Refugee Camp Scenario Calibration of the Adapted PAX Model, Scythe, Issue 6 Workshop 18, 2009
27. Liang, K. H., Wang, K. M. Using simulation and evolutionary algorithms to evaluate the design of mix strategies of decoy and jammers in anti-torpedo tactics, Proceedings of the 2006 Winter Simulation Conference, pp 1299-1306, 2006
28. Yang, A., Abbass, H. A., Sarker, R.. Landscape Dynamics in Multi-agents Simulation Combat Systems, Australian Joint Conference on Artificial Intelligence No. 17, Cairns, pp. 39-50, December 2004
29. Oliehoek, F. A., de Jong, E. D., Vlassis, N. The Parallel Nash Memory for Asymmetric Games, Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, pp. 337-344, 2006
30. Chua, C. L. , Sim, Cpt W. C., Choo, C. S., Tay, V., Automated Red Teaming: An Objective-Based Data Farming Approach for Red Teaming, Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, pp. 1936 - 1942, 2007

31. Easton, J. and Barlow, M. CROCADILE - An Open, Extensible Agent-Based Distillation Engine, *The International Journal of Information and Security*, Volume 8, May 2002, pp. 15-51, 2002
32. CROCADILE home page, <http://www.itee.adfa.edu.au/research/vesl/Croc/index.html> Accessed 7 April 2010
33. Berryman, M. Review of Software Platforms for Agent Based Models, DSTO report DSTO-GD-0532, (Available at <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA485784>, Accessed 7 April 2010), 2005
34. Yang, A., Abbass, H. and Sarker, R. Risk Assessment and Capability Requirements Using WISDOM-II, *Artificial Life and Adaptive Robotics Laboratory Technical Report TR-ARAL-200511017*, 2005
35. Ilachinski, A., Irreducible semi-autonomous adaptive combat (isaac): an artificial life approach to land combat. *Research Memorandum CRM*, pp. 97–61, Alexandria: Center for Naval Analyses, 1997
36. Ilachinski, A., Irreducible semi-autonomous adaptive combat (isaac): An artificial life approach to land combat. *Military Operations Research*, Vol 5, pp. 29–46, 2000.
37. Ilachinski, A., *Artificial War: Multiagent-Based Simulation of Combat*. Singapore: World Scientific Publishing Company, 2004
38. ISAAC/EINSTEIN homepage, www.cna.org/isaac, Accessed 7 April 2010.
39. Lauren, M. and Stephen, R., Mana: Map-aware Nonuniform Automata: A New Zealand approach to scenario modelling. *Journal of Battlefield Technology*, Vol 5, No 1, pp. 27-31, 2002.
40. Galligan, D.P., Anderson M.A. and Lauren, M.K., MANA: Map Aware Non-uniform Automata Version 3.0 Users Manual. Devonport, Auckland, Defence Technology Agency. (Draft version), 2003.
41. Galligan, D.P., Modelling Shared Situational Awareness Using the MANA Model. *Journal of Battlefield Technology* Vol 7, No 3, pp. 35–40, 2004
42. McIntosh, G.C. and Lauren, M.K., Genetic Algorithms Applied to Course-of-Action Development Using the MANA Agent-Based Model, *Journal of Battlefield Technology*, Vol 9, No 3, 2006
43. Straver, M.C., Vincent, E. and Fournier, P., Experiences with the MANA simulation tool. *Defence R&D Canada Valcartier Technical Memorandum TM 2006-404*, August 2006. (Available at <http://pubs.drdc.gc.ca/PDFS/unc51/p525762.pdf>, Accessed 7 April 2010)
44. Seng, C. C., Lian, C.C., L. K. M. Spencer, K.L.M. and Darren, O.S.W., A Co-evolutionary Approach for Military Operational Analysis, in *GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*. New York, NY, USA: ACM, pp. 67–74, 2009.

45. Decraene, J., Hingston, P. and Low, M., Evolving Tactical Plans for Strategy Games using Automated Red Teaming, To be presented at the 2010 World Congress on Computational Intelligence, Barcelona, July 2010.
46. Gioppa, T., Lucas, T. and Sanchez, S., Military Applications of Agent-based Simulations, in Proceedings of the 36th Winter Simulation Conference, pp. 171–180, 2004.
47. Choo, C.S., Chua, C.L., Tay, V., Automated Red Teaming: a Proposed Framework for Military Applications. In: GECCO '07: Proceedings of the 9th annual conference on Genetic and Evolutionary Computation, pp. 1936–1942, ACM, New York, NY, USA, 2007.
48. Chua, C.L., Sim, W.C., Choo, C.S., Tay, V., Automated Red Teaming: An Objective-based Data Farming Approach for Red Teaming. In: Simulation Conference, 2008. WSC 2008. Winter , pp.1456-1462, 7-10 Dec. 2008 <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4736224&isnumber=4736042>, 2008.
49. Xu, Y.L., Low, M., Choo, C.S., Enhancing Automated Red Teaming with Evolvable Simulation. In: GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation, pp. 687–694, ACM, New York, NY, USA, 2009.
50. Lee, M., Ang, D., Hung L.F., Applying Automated Red Teaming in an Urban Ops Scenario. In Scythe 1: Proceedings and Bulletin of the International Data Farming Community, pp. 24–30, Monterey, CA, USA, Naval Postgraduate School, 2006.
51. Sim, W.C., Choo, C.S., M-Tiburcio, F., Lin, K., Shee, M., Applying Automated Red Teaming in a Maritime Scenario. In Scythe 2: Proceedings and Bulletin of the International Data Farming Community, pp. 26-29, Monterey, CA, USA, Naval Postgraduate School, 2006.
52. Wong, A. C. H., Chua, C. L., Lim, Y. K., Kang, S. C., Teo, C. L. J., Lampe, T., Hingston, P., Abbott, B., Applying Automated Red Teaming in a Maritime Scenario. In Scythe 3: Proceedings and Bulletin of the International Data Farming Community, pp. 3-5, Monterey, CA, USA, Naval Postgraduate School, 2007.
53. Torres, G., Buscemi, J., Pickett, K., Hoivik, T., Sanchez, S., Upton, S. and Wan, H., Data Framing in Netcentric Systems Test Planning, In Scythe 6: Proceedings and Bulletin of the International Data Farming Community, pp. 17-20, Monterey, CA, USA, Naval Postgraduate School, 2009.
54. Baez, Major F., Shockley, J. and Aylward, M., Logistics Battle Command Module, , In Scythe 6: Proceedings and Bulletin of the International Data Farming Community, pp. 32-33, Monterey, CA, USA, Naval Postgraduate School, 2009.
55. Henscheid, Z., Middleton, D., Bitinas, E., Pythagoras: An Agent-Based Simulation Environment, Scythe, Issue 1 Workshop 13, 2007.
56. Project Albert Homepage, <http://www.projectalbert.org/>, Accessed 9 April 2010
57. Tisue, S., Wilenski, U., NetLogo: A Simple Environment for Modeling Complexity, International Conference on Complex Systems, Boston, May 16–21, 2004.

58. Schwarz, G. Command and Control in Peace Support Operations Model PAX - Approaching new Challenges in the Modeling of C2, 9th International Command and Control Research and Technology Symposium, 14-16 September, 2004.
59. Lampe, T. A., Schwarz, G. J., Wagner, G., PAX: Designed for Peace Support Operations, Scythe, Issue 2 Workshop 14, 2007.
60. Luke, S., Cioffi-Revilla, C., Panait, L. and Sullivan, K., MASON: A New Multi-Agent Simulation Toolkit. In: Proceedings of the 2004 SwarmFest Workshop. (Available at: <http://cs.gmu.edu/~eclab/projects/mason/publications/SwarmFest04.pdf>, accessed 12 April 2010), 2004.
61. Chi, S., You, Y., Jung, C., Lee, J., and Kim, J., FAMES: fully agent-based modeling & emergent simulation. In *Proceedings of the 2009 Spring Simulation Multiconference* (San Diego, California, March 22 - 27, 2009). Spring Simulation Multiconference. Society for Computer Simulation International, San Diego, CA, 1-8, 2009.
62. Yang, A., Abbass, H.A., Sarker, R., Landscape dynamics in multi-agent simulation combat systems. In: Proceedings of 17th Joint Australian Conference on Artificial Intelligence, LNCS, Cairns, Australia, Springer-Verlag, 2004
63. Yang, A., Abbass, H.A., Sarker, R., WISDOM-II: A network centric model for warfare. In: Ninth International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES 2005), LNCS 3683, Melbourne, Australia, 2005
64. 5. Yang, A., Abbass, H.A., Sarker, R., Curtis, N.J., Evolving capability requirements in WISDOM-II. In Abbass, H.A., Bossamier, T., Wiles, J., eds.: *Advances in Artificial Life, Proceeding of The Second Australian Conference on Artificial Life (ACAL05)*, Sydney, Australia, World Scientific Publisher, pp. 335–348, 2005
65. Yang, A., Abbass, H.A. and Sarker, R., Land Combat Scenario Planning: A Multiobjective Approach, Lecture Notes Computer Science, LNCS vol. 4247, pp. 837-844, 2006.
66. Whitacre, J. M., Abbass, H. A., Sarker, R., Bender, A., and Baker, S. 2008. Strategic positioning in tactical scenario planning. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation* (Atlanta, GA, USA, July 12 - 16, 2008). M. Keijzer, Ed. GECCO '08. ACM, New York, NY, 1081-1088. DOI=<http://doi.acm.org/10.1145/1389095.1389293>, 2008
67. Russel, S. and Norvig, P., Artificial Intelligence, A Modern Approach, 3rd Edition, Prentice Hall, 2010.
68. Yang, A., Abbass, H., Sarker, R. and Barlow, M., Network Centric Multi-Agent Systems: A Novel Architecture, Artificial Life and Adaptive Robotics Laboratory, ALAR Technical Report #TR-ALAR-200504004, University of South Wales, Australia, 2005.
69. Rao, A.S. Georgeff, M.P., BDI agents: From theory to practice. In Proceedings of the 1st International Conference on Multi-Agent Systems, pp. 312–319, San Francisco, USA, 1995.
70. Tidhar, G., Heinze, C., Goss, S., Murray, G., Appl, D. and Lloyd, I., Using intelligent agents in military simulation or “using agents intelligently”. In Proceedings of the Eleventh Innovative Applications of Artificial Intelligence Conference, 1999.

71. North, M.J., Tatara, E., Collier, N.T. and Ozik, J., Visual Agent-based Model Development with Repast Symphony, Proceedings of the Agent 2007 Conference on Complex Interaction and Social Emergence, Argonne National Laboratory, Argonne, IL USA (November 2007). Available at http://www.dis.anl.gov/publications/articles/North_et_al_Repast_Symphony_Tutorial.pdf.
72. Wheeler, S., It Pays to Be Popular: a Study of Civilian Assistance and Guerilla Warfare. Journal of Artificial Societies and Social Simulation 8(4)9, 2005, Available at <http://jasss.soc.surrey.ac.uk/8/4/9.html>.
73. Bosse, T. and Gerritsen, C. Social Simulation and Analysis of the Dynamics of Criminal Hot Spots, Journal of Artificial Societies and Social Simulation 13(2) 5, 2010, Available at <http://jasss.soc.surrey.ac.uk/13/2/5.html>
74. ANTUNES, L., Paolucci, M., and Norton, E. (eds), Multi-Agent-Based Simulation VIII-LNAI 5003. Berlin: Springer, 2008.

Computational Red Teaming: A Review for the Defence Science and Technology Organisation

Cara MacNish

Faculty of Engineering, Computing and Mathematics
The University of Western Australia

Draft, May 2010

1 Introduction — Towards Computational Red Teaming

Red teaming is a process used to help analysts find vulnerabilities in a conceptualised, simulated or rehearsed environment prior to those vulnerabilities being exploited in the real world. It is useful in a wide range of applications, from military operations, to infrastructure and event security, computer network security, and adversarial business practices.

Traditional or *manual red teaming* (MRT) can be a difficult and time consuming process where complex systems and interactions are involved. As human mobility and technology increases, so do the complexity of the systems involved and the potential array of threats. According to Skroch [15] the ground that red teams must cover is growing and systems security is not keeping pace with threats. Skroch argues that because of the impracticality of staffing and funding a sufficient number of red teams to address the problems we face now and into the future, the focus must be on *effectiveness*, and one of the keys to effectiveness is *modelling and simulation* (M&S)

Skroch is quick to stress that M&S cannot replace live (manual) red teaming, and that a live red team has features that cannot be matched by computer simulation. His concern is to determine how simulated red teaming can be used in concert with live red teaming, and in what situations is each more appropriate. Figure 1 shows Skroch's proposed measures of effectiveness for live red teaming versus simulated red teaming, adapted from an earlier study from the University of Wisconsin-Madison [2]. The shaded (green) entries show in which measures he believes there is a distinct advantage of one over the other.

The measures in which Skroch credits a win to M&S are primarily those that appeal to the computational power of simulations and the consequent ability to enumerate solutions over large solution spaces, along with the

Measures	Live Red Team	Red Team M&S
Adaptation, agility, unknowns, creativity	Inherently adaptable within skill set and resources of team.	Immature domain for simulation. Adaptable within limitations of programming.
Breadth of knowledge	Inherently broad range within team or resources available to team.	Limited to that which is provided to and can be effectively used by the simulation.
Fidelity, precision	Inherently broad within skill set and tools available.	Adjustable based upon capability of simulation and ability to model. Potential for fidelity exceeding human abilities.
Learning and unlearning	Inherent ability to learn. Difficult to "unlearn" or forget, thus causing tainting of future red team efforts.	Ability to learn based upon fidelity of model. Ability to "unlearn" or forget what has been done in the past.
Stochastic variation, range of possibilities	Variation can be expensive due to labor costs and limits on time to reproduce the red team effort.	Implicitly possible to create a wide range of variation quickly and at low cost.
Live Virtual Constructive (LVC)	Primarily focused on live. Ability to use tools to engage cyber systems. Inability to truly engage LVC without simulation.	Some systems allow LVC, primarily focus on LV, with few incorporating C. Merging live and constructive red team simulation has potential.
Measureable results, data collection	Often difficult depending on how the red team is instrumented. May slow the red team effort or increase costs.	Implicitly available, all data is usually available and can be recorded with little cost.
Speed, capacity	Usually limited to real time, with exception based upon speculative tools for red team. Limited by number of read team.	Essentially unlimited. Essentially limited only by computing resources available.
Reproducible results	Requires use of methodology and proper selection of team members.	Inherently able to reproduce past events and provide consistent environments for constructive simulations.
Accurate results – ability to be validated and verified, V&V	Requires rigor of process, shadow red team, multiple red teaming, V&V must be reconsidered with each new team	Nature of simulation enables V&V that is sustained across similar simulations or models
Cost	Usually considered higher cost due to labor. May cost less in a small tactical red team effort. Depends on breadth and depth.	Potential to reduce costs for complete coverage of attack spaces, stochastic variations, sensitivity analysis, etc. May have higher cost if only used for a single run.
Time to set up	Depends on size of red team assembled, time to assemble resources for exercise. Scales linearly.	Depends on M&S tool features, architecture, detail of simulation required. Amortizes over number of runs.
System breadth, complexity, entities	Limited by team size, ability to keep data in mind, time and ability to collect data.	Limited by particular M&S system constraints, time and ability to collect data.
Ability to federate	Innate within constraints of communication.	Dependent on particular M&S system. [HLA, DIS, TENA...]
Bias, COI ¹⁸	Depends on individuals, team affiliation.	Depends on affiliation of analysts, programmers.
Efficiency as a cost-benefit value	Based upon quality of process, team composition.	Potential to be more highly efficient than a red team due to automation.
Effectiveness (w/o regard to efficiency)	Depends on domain of use.	Depends on domain of use.
Physical RT [gates, guns, guards, ...]	Mature discipline requires knowledgeable team members, can be highly effective but not necessarily exhaustive.	Physics-based nature of physical systems lends well to M&S. Ability for excellent variable fidelity results and exhaustive results.
Cyber RT [hardware, software, network, ...]	Discipline immature and often driven by team member expertise, access to particular tools, resources	Isolated hacking and fuzzing tools exist with increasing "intelligence" but are primarily scripted. Potential for speed.
Behavioral RT [skill, culture, M&I ¹⁹ , psyop, phishing, ...]	Live humans are the benchmark for behavior in live systems.	Maturing discipline, behavioral modeling is in its infancy. Able to model defined tactics, techniques and procedures with some success.

Figure 1: Measures of effectiveness for manual red teams and red team modelling and simulation, from [15].

recording capacity of computers and reproducibility of results. This is not inconsistent with the traditional view of computer capabilities.

There are two particular measures attributed to human red teaming, however, that would stand out to researchers and practitioners in *Artificial Intelligence* (AI) and *Computational Intelligence* (CI). These areas, and Skroch's assessment of them are:

1. *Adaptation, agility, unknowns, creativity*

Immature domain for simulation. Adaptable within limitations of programming.

2. *Behavioural red teaming [skill, culture, motivation & intent,...]*

Maturing discipline, behavioural modelling is in its infancy. Able to model defined tactics, techniques and procedures with some success.

Skroch is correct to highlight that these are less mature areas in modelling and simulation. As a result they provide a greater challenge for computational systems, but also promise significant benefits. It is largely these areas that we would regard as distinguishing modelling and simulation, which might be regarded as a largely passive implementation of a search tool to aid MRT, from *computational red teaming* (CRT), which employs techniques from AI, CI, adaptive and agent-based systems to generate creative solutions which a human may not have envisaged. This study is primarily focussed on the progress in these areas.

1.1 Computational Red Teaming in a Military Context

This survey paper considers computational red teaming in the context of two specific directions provided by DSTO. The first is the specific focus on CRT as applied to *Improvised Explosive Devices* (IEDs), and models that support the *When*, *Where* and *How* of specific IED events. The second is a specific focus on *fitness landscapes*, a term used in the evolutionary algorithms literature for the surface of solutions mapped out by a *fitness* or *objective function* (see Section 2). We will argue that these two goals do not converge well in the published literature, and that the work on evolutionary approaches is primarily oriented to scenarios that fit well with existing algorithms but do not utilise features necessary for IED analysis. Conversely, in models better oriented to IED red teaming tasks, there is little evidence in the literature showing how evolutionary approaches might be used. We believe this is an important open research problem.

2 Background — Adaptive Algorithms

Both Artificial Intelligence and Computational Intelligence are concerned with instilling in software traits or capabilities traditionally associated with

human behaviour or inspired by the success of biological processes. Examples include reasoning, planning, learning, and adapting to change. These traits and capabilities are often embodied in the context of an *agent* that is able to act autonomously within its environment. An agent will typically perform some kind of goal-based behaviour, which may be derived from declarative information such as its beliefs, desires and intentions, or from numerical information such as maximising some measure of utility. By studying collections of agents, with different communication capabilities and interaction parameters, emergent behaviour of populations may also be observed.

While a number these concepts appear in the literature that follows, there is not room for a comprehensive introduction here. Given the focus on fitness landscapes, however, some further explanation of evolutionary and allied algorithms is warranted.

It can be argued that there are two primary adaptive mechanisms that allow biological organisms to survive and thrive. The first is adaptation during the organism's lifetime, which allows it to deal with changes in its environment. In higher order organisms, this takes place primarily through learning. The second is adaptation through the collective lifetime of species, or *evolution*. The remarkable success of evolution in nature has not gone unnoticed by algorithm developers, who realised that for some problems that are too complex or not well enough understood to write algorithms for directly, it may be possible to *evolve* solutions.

This approach was rapidly adapted to a wide range of problems and can be regarded as a search technique. Examples of these search algorithms include genetic algorithms and evolutionary algorithms or evolutionary computation. They have been found to be particularly effective for "poorly behaved" search spaces where more traditional algorithms may fail. The success of evolutionary algorithms led to proposals for a large number of variants and allied algorithms with similar properties and different strengths and weaknesses. In some cases, such as *particle swarm optimisation*, these were also inspired by nature. In others they simply borrowed key features, such as being based on populations of candidate solutions, with probabilistic mechanisms for generating new candidates (or, equivalently, adapting existing candidates), and an *objective function* or *fitness function* for evaluating the candidates and determining which to propagate to the next generation. This is done in such a way that the best characteristics of the solutions are preserved within future populations, and the performance of the population (ideally) improves. Although the term evolutionary algorithm has come into common use both for the original evolutionary algorithms, as well as the class of algorithms as a whole, I prefer to use the term *adaptive algorithms* for the class of algorithms where possible to avoid the potential misinterpretation that I have fixed on a particular algorithm within this class.

Where a potential solution is parameterised by n real-valued attributes, it can be seen as a point in an n -dimensional space. If the fitness of the solution is included, each solution is represented by a point in $n + 1$ -dimensional space. If the parameters were varied continuously across their ranges in n -dimensions, and the solutions plotted in $n + 1$ -dimensional space, they would form a (possibly discontinuous) hyper surface. This is called the *fitness landscape*. The height of the fitness landscape in the $n + 1$ th dimension indicates the quality of the solution as measured by the fitness function. Normally we seek to find either minima or maxima on this surface. In general there may be many local minima and maxima, and we wish to find the best we can given a certain computing resource. Different algorithms will achieve this to different extents, depending on the shape or characteristics of the surface, or landscape. Studying the fitness landscape is important, not only to inform us about what algorithm(s) might be appropriate, but also to give us some insight or confidence in how well our algorithm is likely to be doing.

It is easy to see why adaptive algorithms are ideally suited to computational red teaming. In the (most common) case of evolving the red team characterisation, for example, each solution, or point in the solution space, represents a different red team approach (different configuration, strategy, characteristics, goals, priorities, etc, depending on the characterisation used). Adaptive algorithms allow us to explore the solution space, taking the best parts of each solution and modifying or combining them to form better ones. To the extent that:

1. the scenario has been adequately modelled
2. the appropriate features (measures, characteristics, attributes, etc) can be extracted, and
3. the algorithm is able to cope with the fitness landscape

an adaptive algorithm will be able to find and zero in on any vulnerabilities in blue's defenses.

Unfortunately, each of these are usually difficult problems! This is particularly true for IED scenarios. They are difficult to model because they involve human decision-making and social factors (it is far easier to model a missile than a human), there may be a wide range of features that may affect behaviour, many of which may not be immediately obvious, and since they involve humans making discrete decisions the fitness landscape is likely to be discontinuous and somewhat erratic.

This paper surveys some of the attempts to date to apply adaptive algorithms to military red teaming, in the context of these problems.

2.1 Adaptive Approaches to CRT

There are four ways in which adaptive algorithms are utilised in red teaming in the literature. The first, as suggested above, is to generate new strategies or improve existing strategies (or parameters that implicitly determine the strategies) of the red team. The idea here is to provide solutions that challenge or defeat the blue team that may not have been foreseen through manual red teaming or “hand tailoring” the software in simulated red teaming.

The second is to help analysts improve the blue team response. Given that red teaming has uncovered vulnerabilities in blue’s defenses, how, if at all, can these “holes” be plugged using existing forces and infrastructure, or is additional help needed?

Following from these two it is easy to see that adaptive algorithms could be used in an iterative, or cyclical, approach. First the red team is adapted to find the most effective adversary against blue’s defensive strategy, then blue’s strategy is adapted to shore up any deficiencies. The red team is then adapted again to search for further or new holes, and so on. This iterative approach, when automated, is called *co-evolution*. Ideally there will be a cumulative improvement in blue’s (and red’s) strategy, and ultimately blue will be shown to be in a stronger position (otherwise it is an indication that additional resources are required).

Finally, adaptive algorithms can be used to adapt the model itself. We will see examples of each of these.

3 Red Team Evolution

3.1 Early Approaches

Early examples of the use of evolution in automated red teaming were proposed by Upton and colleagues [18, 17]. The aim of their approach, which they called *AutoRedTeaming*, was to “use agent-based simulations to simulate proposed security procedures and then allow the Red (threat) agents to evolve capabilities. . . with the goal of discovering means to thwart, evade, or otherwise exploit gaps in Blue’s security procedures. . .” [17].

Their approach is typical of much of the work that followed. They use simple reactive agents whose capabilities and hence behaviour is controlled by a set of parameters. The blue agents’ parameters are fixed, while the red agents’ parameters are evolved using an evolutionary algorithm. Their approach is trialled in a case study in which they developed an agent-based scenario involving the defense of a fixed structure, and evolved the parameters using an evolutionary programming algorithm with simple mutation and tournament selection.

While no conclusive results are reported, they do raise one of the key

issues that we focus on throughout this study, simulation *expressiveness*, and the limited flexibility of generating solutions using only parameter setting of the agents. They also propose the concept of *evolvable simulations*, claiming that searching for not just parameter settings but changes in the simulation agent’s structure will “allow for much more robust and powerful AutoRedTeaming”. This is a topic we return to in Section 6.

3.2 The ART Framework

Following from the work of Upton *et al*, Choo *et al* [3] propose a modular framework for red teaming called *Automated Red Teaming* (ART). Again the goal is to use evolutionary algorithms to improve red team performance, with the ultimate goal of reducing surprises and improving the robustness of blue operations. The authors argue that this is necessary since the vulnerability assessments made by human experts are bounded by their knowledge of the subject matter, although there is little evidence in the paper to suggest that the proposed techniques address knowledge bounds as opposed to search capacity.

Choo *et al* argue that Upton *et al*’s work has two shortcomings. First, it utilises evolutionary programming with a single objective function, and that this does not model real-world situations where multiple criteria for success might exist. This is a curious distinction on which to base their work. The authors refer to studies comparing and ranking various multi-objective evolutionary algorithms, and ultimately select *Strength Pareto Evolutionary Algorithm Version 2* (SPEA2), in part because it is considered to be more robust when dealing with high-dimensional objective spaces (which their experiments are not). They then proceed to demonstrate their system on a scenario with a single objective function.

The second stated shortcoming is the potential difficulty of altering legacy simulation models for use with red teaming. The ART framework addresses this by using a modular structure which allows different simulations to be “plugged in”. Before this can happen, however, wrappers need to be written to obtain lists of parameters from the simulation, as well as input and output wrappers. It is not clear how much work is required for this compared with Upton *et al*’s system.

Nevertheless, Choo *et al*’s work does provide a proof of concept for a modularised automated red teaming system working on a simplified “real world” problem. The framework consists of four parts, the ART backbone modules, modules for wrapping the simulation model, a module that schedules the submission of jobs to clusters and collects their outputs, and a module containing the evolutionary algorithm.

The case study provided is an urban ops scenario in which the blue team must raid and capture an installation defended by the red team, in the presence of hostile civilians. It is modelled using the *Map Aware Non-*

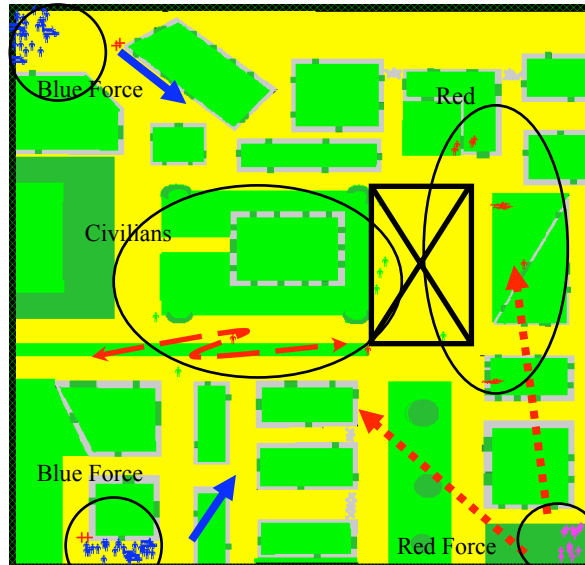


Figure 2: An urban ops scenario modelled in the MANA simulation tool, from Chua *et al* [3].

uniform Automata (MANA) agent-based simulation tool developed by the Defence Technology Agency in New Zealand [8], and illustrated in Figure 2. The blue team consists of 2 platoons of 21 soldiers along with 3 soft-skin vehicles. The red team consists of 7 soldiers positioned in the proximity of the installation, as well as 4 snipers and 14 additional soldiers that can be called upon when outnumbered.

Red team activity is controlled by a set of parameters that determine behaviour such as tendency to cluster, aggressiveness, cohesion and stealth. These parameters are adapted by the evolutionary algorithm (SPEA2), which was run through 40 iterations with 30 children and 30 parents. The results, reproduced in Figure 3, show a substantial reduction in red attrition, along with a more modest increase in blue attrition.

The mean final values of the parameters are used to suggest the characteristics of an effective red force. For example, the force should be highly stealthy, slightly aggressive and tend not to move cohesively. The variance of a parameter is interpreted as a measure of whether that characteristic has a significant impact on performance.

While these results are encouraging, further research is needed to fully interpret them. Without any analysis of the fitness landscape, it is unclear whether an evolutionary algorithm is in fact the best way to search for solutions. It is possible, for example, that a simple hill climb could produce superior results with the same computing resource. The landscape could

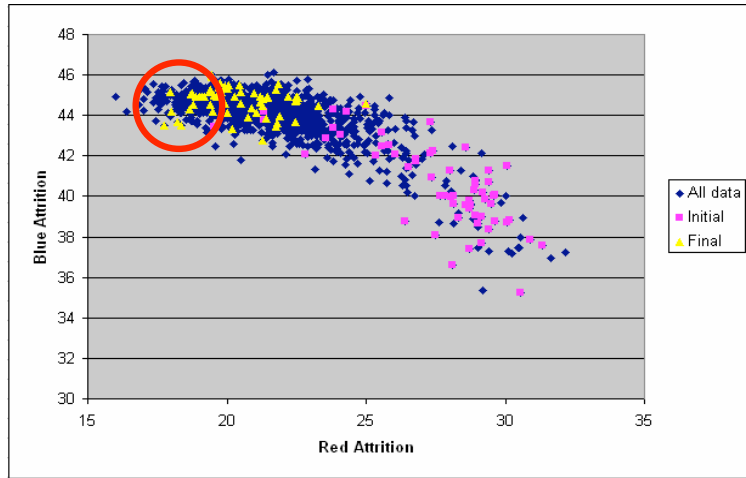


Figure 3: Red versus blue infantry attrition, from Choo *et al* [3].

also affect the variance of particular parameters with, for example, a higher variance indicative of multiple optima rather than insignificance of a parameter.

As Figure 3 shows, there is also little evidence to suggest that the 40 iterations is sufficient for the population to converge, nor that if it does converge, it will do so on a global optimum. Furthermore, the fact that the best solutions (on the top left) are not the final solutions (in yellow) suggests that greater elitism, better tuning, or an alternative algorithm may be required to avoid losing optima.

3.3 Comparison of Computational and Manual Red Teaming

In subsequent work Chua and colleagues seek to validate their work by comparing the findings of ART with those obtained through manual red teaming [4]. Their paper describes the findings of exercises conducted at two International Data Farming Workshops, led by the US Naval Postgraduate School’s Simulation Experiments and Efficient Designs Center for Data Farming. The authors present ART as a specialised variant of data farming, where “instead of exploring the entire parameter space to generate a response surface, a point (or objective) on the response surface is first identified, followed by a search for the parameters that result in this point”. Once again, this is somewhat in contrast to their framing of the task as a multi-objective optimisation problem. Nevertheless, the study represents an interesting attempt to verify the element of surprise that can be obtained by evolutionary techniques, when compared with human creativity.

The study makes use of the ART framework described in Section 3.2,

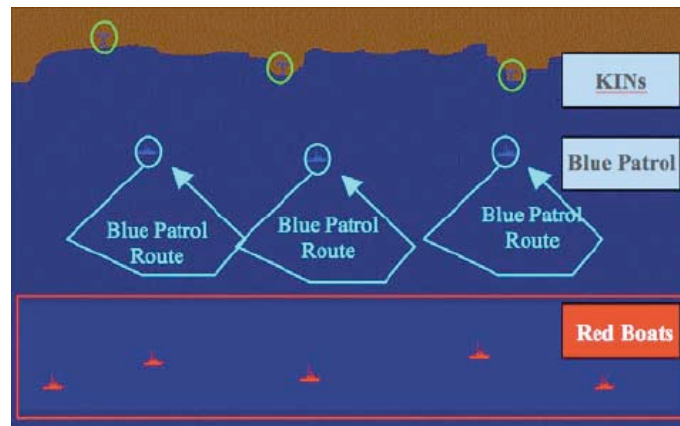


Figure 4: A maritime key installation protection scenario in MANA, from Chua *et al* [4].

with SPEA2 again used as the search engine, and MANA used as the simulation engine. The paper presents two case studies, the first involving the protection of key installations, and the second involving protection of an anchorage.

The first scenario consists of three coastal surveillance radar installations, protected by three blue patrol boats, as illustrated in Figure 4. The radar installations have minimal machine gun protection, and are able to communicate detected targets to the patrol boats. The red force consists of five red boats modelled as small fishing boats which act independently, simulating a peace-time threat. (In fact while the red boats are not given the ability to communicate, the tactics generated suggest coordination in advance of an attack rather than complete independence.) The fishing boats are loaded with explosives, and might be regarded as a maritime example of IEDs, although their attacks are not well concealed.

The scenario is described with two objectives (or *measures of effectiveness*). The first is the “mean red mission success”. Curiously a mission was apparently considered successful if at least one boat managed to penetrate the blue defence, irrespective of whether it inflicted damage on a key installation, and perhaps for this reason, combined with red outnumbering blue, mission success is 100% in all trials. The second objective is to minimise mean red attrition.

The manual team developed two plans, one in which the red boats moved down the flanks, and one which combined flanking with a saturated attack in the centre, with mean red attrition of 0.85 and 3.05 respectively.

Automated red teaming used SPEA2, with aggression, cohesiveness and determination as the adapted parameters. The paper does not comprehen-

sively describe the experimental setup, for example it does not specify how the initial conditions are determined or how the boat directions are set or evolved. The automated approach produced a “decoy” tactic that the authors note surprised the team, where one of the small boats lured a blue boat away to leave an opening for two further red boats. This produced a red attrition of 1.89. Although this did not prove to be the best tactic, it did demonstrate the ability of the ART approach to simulate “lateral thinking” by coming up with a solution that the manual team had not considered.

The second scenario consists of an anchorage protection task, with blue patrols around an anchorage containing commercial vessels. The goal of the red team, again consisting of five small boats loaded with explosives, is to penetrate the blue patrols and inflict damage on the neutral vessels. In this case there are three objectives listed: mean red mission success, mean red attrition, and mean neutral shipping destroyed.

The manual team developed a single strategy that involved a saturated attack on the anchorage area. This resulted in 100% mean red mission success, mean red attrition of 1.96, and 3.05 for neutral shipping destroyed.

The ART approach was again apparently achieved through adaptation of aggression, cohesiveness and determination. It arrived at an improved plan with mean red mission success of 100%, a lower mean red attrition of 0.48, and a higher mean neutral shipping destroyed of 4.52.

There are many questions left unanswered in this case study however. Again it is not clear how the initial conditions are set. ART is said to have generated a similar plan to the manual saturation tactic, consisting of a “simultaneous red attack towards the centre of the anchorage area with re-attack flexibilities”. It is not clear what “re-attack flexibilities” means in this context, or more importantly how they can be evolved using only the parameters described. The authors go on to claim that “the ART results pushed it further with red boats traversing the anchorage looking for dispersed vessels. The ART design has given insights that it is not enough to just stop the red boats from reaching the anchorage but it is also important to keep the leaked boats from maneuvering within the anchorage”. While these may be reasonable intuitive interpretations of the routes generated by the evolutionary process, it is not possible to generate these routes with the parameters stated alone. It is also not clear whether this degree of freedom was offered to the human team.

This leads to the question how the manual team were instructed and the extent to which they were constrained. The results summary provided lists the tactics generated by the manual team in terms of the aggression, cohesiveness and determination parameter settings. Since it would be difficult to “reverse engineer” these values from the resulting behaviour, this suggests that rather than “driving” (or setting route plans) for the red boats, the manual team may have been restricted to manipulating the parameters in an attempt to achieve the outcome they wanted — effectively the same task

that the automated system is seeking to achieve. If this is the case then you would expect the human team to have little chance of outperforming the automated system, given the search capacity of the automated system. The human team would need to appeal to their own strengths in pattern analysis and spatio-temporal reasoning.

Finally, it is not clear how the manual team or the automated approach dealt with the multiple objectives. The motivation for the use of a multi-objective evolutionary algorithm such as SPEA2 is to generate a pareto-optimal front of solutions, each of which is better than the others in at least one objective. This allows a human to investigate alternative solutions that are better in some aspects than others. There is no evidence of this in the paper, where only a single solution is presented from ART. Furthermore, if the manual team is required to search for solutions by manipulating the parameters, it would require considerable work on their part to generate this multi-objective front and compare solutions.

4 Blue Team Evolution and Fitness Landscapes

Yang *et al* [23] take a different approach to evolutionary red teaming. In this approach evolution is used more directly as a “combat enhancement technology” to develop strategic and tactical responses for the blue team in response to different red team strategies or characteristics. The authors highlight the cost of the evaluations required to search the space of potential solutions. This is due both to the inherent cost of playing out simulations for each fitness evaluation, and the added complexity afforded by the stochastic nature of a simulation. This leads to a focus on trying to understand the fitness landscape in which solutions are sought.

In order to carry out their analysis of fitness landscapes, Yang *et al* utilise a multi-agent land combat simulation system known as the *Warfare Intelligent System for Dynamic Optimisation of Missions* (WISDOM) developed at the University of New South Wales at the Australian Defence Force Academy. WISDOM permits investigation of emergent team behaviour as a result of individual agent characteristics. These characteristics include sensors, capabilities, movements, communications and health. Of most interest in Yang *et al*’s study is an agent’s movements, which are in turn determined by a set of parameters defining its “personality”. This consists of five categories of weights that determine the agent’s desire to move toward a healthy or injured friend, a healthy or injured opponent, and the target (a flag). The first four categories consist of two weights that are combined with vision and communication information. An additional weight, the “probability to hit”, reflects the agent’s firing skills. The result is a vector of ten real-valued weights that determine an agent’s behaviour, and can be used as decision variables (or parameters) in an adaptive algorithm.

Two random decisions are identified in the movement algorithm which result in the stochastic nature of simulations. Firstly, movements between locations (cells) are determined by a penalty function based on the above weights. Where that results in a tie, one of the winning moves is chosen at random. Secondly, an agent always fires at the nearest enemy. Again an enemy is randomly selected in the case of ties.

It is perhaps worth noting at this point that this kind of nondeterminism is, at least at first sight, very different to the nondeterminism that plays a key role in “surprise” and must be modelled in the context of asymmetric warfare and IEDs. It could be argued that this nondeterminism is an artifact of the discretised nature of the model and would not occur in a continuous model where the probability of equal penalties approaches zero.

Yang *et al*’s work examines the nature of the search space, or fitness landscape, for blue team solutions. To achieve this they pre-define six red team strategies, with varying degrees of aggressiveness, tendency to cluster, and goal orientation, encoded in the agents’ personality characteristics. An objective function is defined that maximises blue team health at the end of a simulation, and this is used to derive two fitness functions. The first averages the objective function over 100 simulations, the second normalises the average against the standard deviation.

Two experiments are carried out. The first examines the fitness landscape by taking 10 000 step random walks with Gaussian deviations, and examining the time series as advocated by Vassilev *et al*’s “information content” analysis [19]. The authors define *signal worst* to be the difference between the best and worst fitness values encountered during the search, and use this to calculate the “information stability”, “information content” and “partial information content”, as well as an expected number of optima. This allows the authors to conjecture differences in the nature of the landscapes corresponding to different red team strategies. For example, the result which stands out the most is the information stability for the aggressive and very aggressive strategies with normalised average fitness, which is an order of magnitude lower than the other results. These strategies also have somewhat higher partial information content, leading to the conclusion that these strategies lead to a highly multimodal normalised fitness function. The information content, on the other hand, is similar across the strategies suggesting a similar degree of ruggedness.

The second experiment seeks to test the findings of the landscape analysis using a simple $(1 + 1)$ evolutionary strategy or stochastic hill-climber. While there are certainly similar characteristics between the random walk and hill-climber results, in particular the distinctly different normalised average fitness results for aggressive and very aggressive strategies, it is less clear how the information content analysis manifests in the hill climber behaviour. For example, non-aggressive strategies got “stuck” in an attractor with normalised fitness value 200, despite the landscape analysis predict-

ing a higher number of peaks for the aggressive strategies. Thus, while the fitness scores from the trials are useful in conjecturing a number of other intuitive results, it is not clear how much benefit is afforded by the information content approach.

4.1 Fitness Landscape Analysis and Sensitivity

An important question about this kind of fine-grained landscape analysis is its robustness in the light of uncertainty about the fidelity of the model. Yang *et al* touch on this issue by raising the question of sensitivity to personality parameters. To address this they conduct sensitivity tests in which some of the weights are perturbed by a small step of 0.05, and note minimal changes to the fitness landscape and information theoretic measures.

There are many other factors involved in the model and the initial conditions, however, that may far outweigh a small perturbation of the weights, and even potentially larger distinctions such as the differences in red team aggression. These range from the choice of model, characteristics and capabilities of the agents, relative weight ranges, the discretisation of moves, the way in which the various factors are apportioned in the movement penalty function, and the terrain, numbers and positions of agents in the experimental scenarios.

The properties of the fitness function will affect the search algorithm in two ways. First, they may effect the speed at which good solutions are found. It seems likely that in the military scenario this level of fine tuning is less important than the quality of solutions found. Secondly, in more extreme cases, the landscape properties may result in some kinds of adaptive algorithms never reaching solutions that others may find, for example as a result of premature convergence. However the exact relationship between landscape properties and adaptive algorithm choice remains unclear.

For these reasons, while a thorough examination of the fitness landscape is certainly of interest from a research perspective, this degree of fine tuning may be regarded as secondary to the appropriateness of the model.

5 Co-evolution

Co-evolution is an approach that could potentially be used for both purposes of generating good red team strategies to challenge a (manual) blue team's response, or generating blue strategies to combat an increasingly capable red team. Choo *et al* [14] seek to use co-evolution for the latter, stating that the "overall goal is to complement the manually intensive actions-reactions process in developing (automatically) a blue plan that performs well and is relatively robust even in the face of an adaptive red adversary".

Choo *et al* describe an architecture called *Automated Co-evolution* (ACE) which facilitates the use of a competitive co-evolution algorithm (CCEA) on

a military scenario simulator, farming the evaluations to a parallel cluster. The CCEA is characterised as a host and parasite relationship, although in this application the red and blue teams share equal status. In each iteration, the members of one team are evolved against scenarios where the strategy of the other team is fixed, and then the second team is similarly evolved against the first. In this paper an “All vs Best” approach is used, where all red team solutions are evolved against a single scenario involving the best blue team solution, and vice versa.

The ACE framework is designed to allow different algorithms and simulation engines to be “plugged in”. The simulation engine used in the paper is MANA [8].

The paper curiously considers four adaptive algorithms, one of which — ant colony optimisation — is not appropriate for the task and another of which — SPEA2 — is designed for multi-objective optimisation. Having discarded these two, a particle swarm optimiser (PSO) and *Elite Pareto Genetic Algorithm* (EPGA), are implemented.

A case study is presented using a maritime anchorage protection scenario in MANA similar to the second case study in [4] covered in Section 3.3. Unlike [4], however, this paper explicitly includes initial conditions and way points in the adapted parameters, showing how vessel routes are constructed.

The case study involves neutral ships within a central rectangular area, three blue patrol vehicles parameterised by a home position and two way points (thus generating triangular patrol routes) and five red craft parameterised by start and end points and an intermediate way point (thus generating on attack trajectories with a single change of direction). The objective function incorporates the number of neutral ships destroyed and the number of red ships destroyed.

The study shows that on each co-evolution cycle, the blue team gains significant ascendancy during its evolutionary period, then the red team gains significant ascendancy during its evolutionary period. This “see saw” effect is explained in terms of red exploiting defense gaps, and blue patching them. However it is hardly surprising, since it is clearly an underconstrained problem, where there is no strategy that blue can adopt which cannot be penetrated by red with a reasonable search. Similarly, for any fixed red attack, blue is able to search for a path that blocks it.

What appears to be missing in these results is the very thing that co-evolution promises to provide — a *cumulative* improvement in strategies. This would be indicated by red’s solutions resulting in lower neutral ship attrition over time. The results show that red does indeed achieve less neutral attrition in some co-evolution cycles, but returns to high attrition. As such it fails to satisfy the original goal of developing a blue plan that performs well *and is relatively robust*. Each plan appears to be roughly as vulnerable as the last.

There are many possible reasons for this lack of cumulative performance

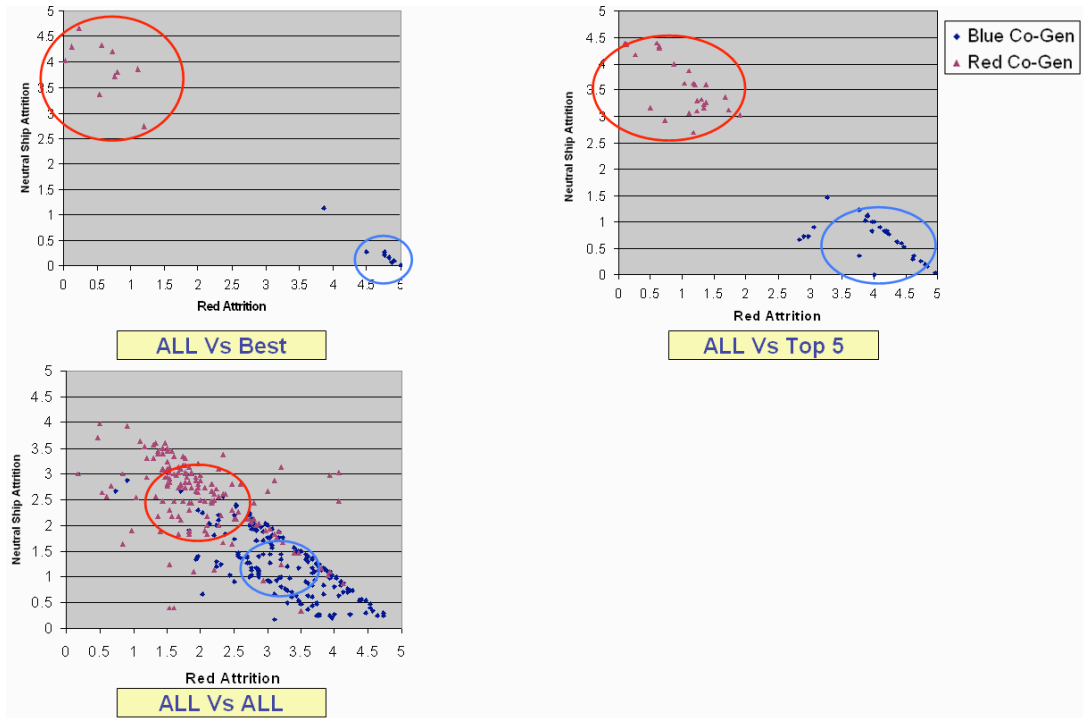


Figure 5: Neutral ship versus red attrition for different evaluation strategies, from Choo *et al* [14].

that do not rule out the efficacy of co-evolution as an approach. One explanation is that the problem is so underconstrained that there simply are no more robust strategies available to blue. However other comparisons in the paper suggest this is not the case.

A likely explanation is that the co-evolution approach is over-fitting. This again would not be surprising, since an “All vs Best” evaluation strategy is used, for the purposes of computational efficiency, in each co-evolution step. This means that blue is adapting solely to perform against the best red attack. Recognising this, the authors also provide a comparison of different evaluation strategies, reproduced in Figure 5. The robustness of blue is indicated by the height of the red markers on the vertical (neutral attrition) axis. It can be seen that the more of the population that is used in evaluations (from Best, to Top 5, to All), the more the red markers are dragged to the middle (with a similar trade-off for blue). This is because blue (and red) is forced to compromise specificity for generality. This could be enforced further by other methods such as using a different adaptive algorithm, as illustrated by Choo *et al*’s comparison of EPGA and PSO (although the

authors argue the more robust algorithm, PSO, is inferior), and reducing the number of EA generations in each co-evolution cycle.

6 Evolving Models

Given the difficulties of appropriately modeling red teaming scenarios, the work of Xu *et al* on *Enhancing Automated Red Teaming with Evolvable Simulation* [22] appears enticing. The paper sets out to investigate “the benefits of Evolvable Simulation, which involves evolution of the structure of a simulation model”. The paper uses the MANA simulator [8] embedded within the ART framework [3].

The work makes use of a maritime case study simulating the protection of commercial ships within an anchorage, similar to the anchorage protection scenario in Chua *et al* [4]. In this case however there are two groups of blue patrol ships with different speed and firing accuracy capabilities and a fixed patrolling strategy. The red ships are required to find a route through the patrol boats and back to safety, inflicting a degree of damage on the way according to a “hit chance”.

In order to search for routes the red team makes use of a *Time-Variant Multi-Objective Particle Swarm Optimisation* (TVMOPSO) algorithm [16]. It is again unclear why this algorithm is chosen, particularly as a single objective is stated (the “mean commercial vessels casualty”). The decision variables include aggressiveness, cohesiveness and determination, in addition to the start and final positions and a number of intermediate waypoints.

Sadly the paper delivers little of its promise to investigate evolving the structure of the simulation model. It turns out the evolution of the model’s structure is limited to altering “the size of the search space without affecting its basic composition such as the number of red and blue ships and their respective hit chances”. The “evolution” is done manually, and is limited to changing the number of way points used for red craft. This is perhaps better regarded as a path adaptation that allows varying length, rather than a model adaptation.

The size of the search space increases exponentially with the number of way points. The results show, not surprisingly, that with sufficient evaluations available to the TVMOPSO algorithm, increasing the number of way points produces better performance from the red antagonists. The red ships have more degrees of freedom in selecting approach paths. However as the number of way points increases further, the performance deteriorates. In this case the number of evaluations permitted (fixed at 5000 and 10000) has become insufficient to find good solutions in the larger search space.

The paper concludes that “the insights obtained in this work show that ES [evolvable simulation] is a useful methodology which allows the decision makers to enhance their understanding [of] military operational tactics”.

It could be argued that rather it is a demonstration of the relationship between degrees of freedom (and hence search space size) and quality of solution found, for a fixed number of evaluations.

7 Alternative Models and Modelling Frameworks

The literature on evolutionary approaches to computational red teaming demonstrates the basic principle of using adaptive techniques to generate alternative behaviours and strategies for the red and blue teams that outperform existing strategies. There is a pattern however in the models and scenarios used for the case studies and demonstrations, which tend to focus on adapting simple continuous real-valued “game-engine-like” properties such as cohesiveness, accuracy, strength and waypoint locations. This is no doubt because, returning to our criteria from Section 2, these kinds of scenarios are easier to model (agents in the models can be moved in each iteration using simple numerical functions that weight the real-valued parameters), it is easier to extract and adapt suitable features, and it is likely that the fitness landscapes are smoother and easier to navigate.

These kinds of models however don’t appear well suited to IED scenarios, which are considerably more complex. It seems likely that suitable models will be required to incorporate cognitive and social aspects, rather than focus on “reactive” movement.

In the remainder of the paper we review some models and modelling frameworks that, while they may not have appeared in the literature for use with either adaptive algorithms or red teaming, provide some pointers for suitable model development for future work.

7.1 Hybrid Cognitive Architectures

Cil and Mala [5] present an architecture for modelling and simulation in asymmetric warfare. They introduce their work by focussing on the fundamental role played by uncertainty in warfare. “Uncertainty, as a function of asymmetry, has increased with the spread of technology and the juxtaposition of conflicting aims, not only between nation-states, but also between non-state actors. As the potential for asymmetry increases, so do the level of uncertainty and the potential for tactical, operational, and strategic surprise. Asymmetry is really nothing more than taking the level of uncertainty, or surprise, to a new level that involves novel ways, means, or even ends.” This focus suggests some promise for scenario modeling involving IEDs.

Cil and Mala’s system is motivated by citing a number of limitations of existing agent-based models for simulating and studying complex adaptive systems (CAS). These include:

- The difficulty of validation [24].

- The lack of clarity between agent-centric and organizational-centric methods [20]. Cil and Mala claim that existing multi-agent systems either focus on modeling individual agents with limited support for interactions, or focus on the agent society and limit autonomous behaviours of single agents.
- The lack of an explicit and auditable model of interaction by combining agents and their interactions in a single model.
- The trade off between *cognitive* and *reactive* agent systems — cognitive systems are able to reason about actions but quickly become intractable for large numbers of agents, while reactive systems cope with scale better but it is difficult to understand or validate their behaviour since there is no reasoning.

These issues, and particularly the last, provide motivation for their proposed “two layer hybrid architecture” which, they claim, matches the needs of future multi-dimensional warfare. They further claim that neither cognitive nor reactive agent architectures alone are capable of solving real world problems.

The proposed *Multi-agent based asymmetric combat simulation architecture* (ACOMSIM) loosely follows the hierarchical structure popularised by Wooldridge and Jennings [21], in which the top layer is responsible for planning and decision-making, and the bottom layer reacts to simulations from the environment. The top layer of ACOMSIM consists of seven “cognitive” agents and six associated databases. The agents are a Mission Analysis Agent, a Mission Time Scheduling Agent, an Enemy Situation Analysing Agent, an Own Situation Analysing Agent, a Terrain Analysing Agent, a Logistic Agent, and an Action Generating Agent. The databases contain information on intelligence; environment (weather); terrain; tactics, techniques and procedures; enemy tactics, techniques and procedures; and logistics. The architecture is illustrated in Figure 6.

The second layer uses MANA [8] as a simulation model to evaluate the top level plans and provide feedback to the commander. The authors claim the system is distinctive in planning, executing in a virtual environment, and providing detailed results all within an agent-based architecture. The architecture is also claimed to have strengths in scalability, heterogeneity, rationality, adaptivity, sociality, an explicit model of interaction, and ability to reason about emergent behaviours.

The hybrid architecture appears to offer significant potential for generating, and evolving, a much broader range of red teaming scenarios than can be offered by reactive simulations alone. While the latter tend to focus on low-level features such as personality characteristics, geographic or topological positioning, and constraints and freedom of movement, the hybrid framework should allow access to higher-level, symbolically represented

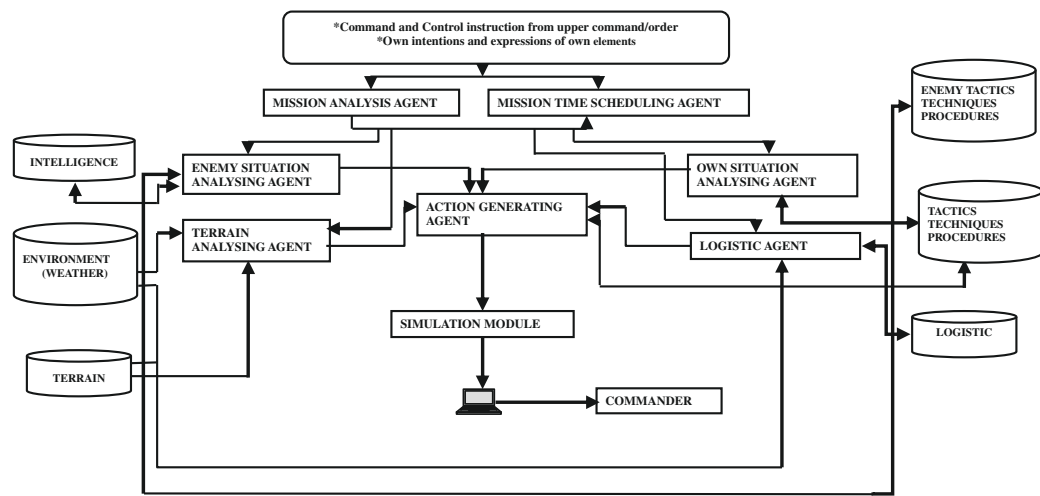


Figure 6: The ACOMSIM architecture, from [5].

features that directly affect high-level planning processes.

7.1.1 Red Team Evolution

Two components in the ACOMSIM framework bear most promise for setting up asymmetric warfare scenarios, particularly of the kind involving IEDs, as part of a computational red teaming exercise. These are:

Enemy Tactics, Techniques and Procedures Database

The primary purpose of this database is to predict enemy actions using stored tactics, techniques, and operational procedures of enemy forces. These are primarily retrieved from past intelligence. The intention of this database is to predict enemy actions using stored data so that appropriate precautions can be taken.

This (in conjunction with the Terrain and Environment Databases) is potentially where adaptive algorithms could be used for setting up alternative red teaming scenarios. Just as a real enemy would change their tactics and techniques over time in an attempt to achieve greater success, adaptive algorithms could be used to preempt such changes. This may be considered the primary vehicle for setting up the *How* of IED scenarios.

Intelligence Database

This database stores information about enemy actions, capabilities, locations, activities and leaders, retrieved from human, technical and

signal intelligence, including air vehicles and satellites. It may be updated in real time.

This is the primary location where adaptive algorithms could be used to test the red team response to changing, real-time intelligence and its implications for red team survival. While the tactics and techniques database may reflect change “in the large” and remain constant during the period in which a scenario is played out, the intelligence database may reflect real-time changes or changes “in the small” that are critical to asymmetric warfare.

This is where “free will” of the opposing forces manifests. While the enemy may use standard procedures and tactics that can be predicted with a degree of statistical accuracy from past behaviour, it is surprising deviations from this that may produce the most danger. In a highly asymmetric situation, the enemy will rely on deviating from expected behaviour to achieve surprise, and any clues from intelligence may be vital. This may be considered the primary vehicle for setting up the *When* and *Where* of IED scenarios.

7.1.2 Blue Team Evolution

Complementary to the Enemy Tactics, Techniques and Procedures Database is a Tactics, Techniques and Procedures Database that is used to find the best operational plan for the friendly forces. This feeds into the Own Situation Analysing Agent and the Action Generating Agent. It is potentially among these components that blue team evolution would take place.

7.1.3 Issues

Unfortunately there is insufficient detail in Cil and Mala’s paper to be able to conjecture what form the characteristic vectors would take in an adaptive or evolutionary approach. The enemy tactics and techniques database is illustrated by a simple UML diagram of an enemy force structure. It is difficult to make any link between this and the stated knowledge of tactics, techniques and procedures claimed for the database, of its use in predicting enemy actions. Similarly, the description of the Intelligence Database contains no details about how the information is represented, or passed to the Enemy Situation Analysing Agent.

At best, then, Cil and Mala’s paper can be considered a guide to an architecture that would support the kind of data needed to apply adaptive algorithms to red teaming with IED scenarios.

7.2 Social Networks and Agent Development Toolkits

An important aspect of understanding IED threats is understanding, and modelling, the human networks behind them. A number of agent-based

development tools have been developed for modelling and understanding human and social networks, building on the social complexity simulator SWARM [11]. One of the more promising and easily extendable toolkits, which may provide a good basis for further development work, is MASON [9, 10].

MASON is described as a “fast, easily extendable, discrete-event, multi-agent simulation toolkit” written in Java. It is an open-source project out of George Mason University. While it was designed particularly with large scale swam-style multiagent systems in mind, it was intended to be flexible enough to serve a range of research areas from robotics and machine learning to modeling social systems and networks. To achieve this it uses a minimal generic core that is readily extendible to domain specific tasks.

A key principle of MASON is a strict adherence to separation of model and view. The independence of the model from the view aids efficiency (for example ease of farming to clusters), robustness (for example checkpointing and resumption on power failure) and flexibility (changing platforms, views, and so on). This is fundamental to the system’s design goals of [10]:

- A small, fast, easily understood, and easily modified core.
- Separate, extensible visualisation in 2D and 3D.
- Production of identical results independent of platform.
- Checkpointing any model to disk such that it can be resumed on any platform with or without visualisation.
- Efficient support of up to a million agents without visualisation.
- Efficient support for as many agents as possible under visualisation (limited by memory).
- Easy embedding into larger existing libraries, including having multiple instantiations of the system coexisting in memory.

When wrapped with a console (as part of the visualisation layer) or batch controller such as the evolutionary computation system ECJ (a sister project from George Mason) it could be regarded as implementing the classic Model-View-Controller (MVC) design pattern.

The model layer in MASON contains agents (or “steppable” objects) along with a discrete-event schedule for triggering agent actions, and *fields* that relate objects to locations (in a notional space). The schedule provides facilities for regular checkpointing. Fields are provided for arrays, grids, continuous space and networks or graphs.

The visualisation layer maintains a number of *displays* that provide 2D or 3D views of fields in the model. The displays hold *field portrayals*, responsible for drawing fields, which in turn associate *simple portrayals* with

objects in the field. The simple portrayals can in turn call *inspectors* to inspect or modify object parameters in response to user requests.

It is intended that extensions to MASON will be provided, or contributed, for more specific problem domains. Some examples of extensions that are available include the evolutionary computation system ECJ, social network analysis, and rigid body 2D physics modelling. It is also possible to connect MASON to a range of external packages.

MASON's design borrows from other multiagent simulators, in particular the mobile robotics simulators TeamBots [1] and Player/Stage [7], and the social complexity simulators SWARM [11], Ascape [13] and RePast [12]. However Luke *et al* [10] claim that MASON is distinctive in its combination of extensibility, small footprint, speed in both the underlying model and visualisation, and most importantly in its separation of model and visualisation and its ability to link and unlink these dynamically.

Example applications of MASON include network intrusion, urban traffic simulation and cooperative target observation in unmanned aerial vehicles, among many others.

7.3 Social Networking and Military Risk Assessment

Building on the the social networking approach, Crossman *et al* [6] provide a significant attempt at combining agent-based social networking models with the tasks of planning and risk assessment in a military context, including IED threats. They argue that while descriptive social modelling tools are valuable for understanding organisations and their structures, or *situational awareness* (SA), they provide little help to analysts in projecting future activities. Following collection of intelligence, SA is just the first step in an analysis process or cycle that includes *projection* of future activities, *planning* a response, and collecting further evidence. To help achieve this they are developing a generative social network, physical process and environmental model to help analysts project future activities in space, time, and socio-political dimensions.

The authors concede that projecting future activities of complex networks of humans ranges from very difficult to nearly impossible as the grain size of the analysis decreases from aggregate groups and wider time windows to specific people, times and events. Nevertheless, they argue, analysts are required to make these judgements, and do so in a somewhat idiosyncratic or *ad hoc* way with little help to ensure that all the bases are covered. For example, analysts may not explore all possibilities, or may miss second or third order effects, longer term effects, or wider regional effects.

The authors are also careful to address the purpose and context of an automated system. They concede that because of the complexity and uncertainty in making predictions that involve people and their decisions, some believe that building such systems is not possible or may give misleading

results. To address this they make a clear distinction between systems of *predictive analysis*, which may be used to “replace thinking”, and systems for *anticipatory analysis*, which should be used to “aid thinking”.

The paper proceeds to describe a system in development called *Defeating Enemy Forces United to Strike with Explosives* (DEFUSE) for anticipatory analysis in the IED domain. The goal of the system is to “rapidly assess future risks, explore contingencies, and plan efficient (intelligence) collection that can confirm or deny hypotheses about what might occur.” The intended output of the system consists of a time sequence of projected activities as well as a collection of named areas of interest.

Like ACOMSIM, DEFUSE is a layered architecture, consisting of a *dynamic network layer* (DNL), *process layer* (PL) and *geospatial layer*. The dynamic network layer consists of a network of *actors*, representing for example leaders, key people and populations, and their relationships. Each actor is associated with a dynamic agent which acts in order to achieve its goals, or increase its *expected utility*. Agents make decisions on the basis of their *beliefs*, *desires* and *intentions*, along with rules encoding the way activity changes the state of the world. Agents beliefs and goals do not have to be consistent, and the authors argue that this has the advantage that “it provides a computational mechanism for generating major shifts (or tipping points)” that can lead to dramatic shifts in agents’ behaviours.

The process layer ties the dynamic and geospatial layers together, and provides a way of modelling physical processes that are not necessary to model in spatial detail (such as building an IED). The PL also models the resources necessary to execute tasks.

The geospatial layer contains terrain databases, and “terrain overlays” — fields which represent distributions of environment properties such as threats, and create the virtual landscape over which entities make decisions and move. Entities representing individuals, such as bomb makers, and teams, such as deployment teams, are represented as “polyagents” which consist of an avatar and a set of “ghosts”. The ghosts represent various possible projections of the agents, with different characteristics, into the future in the virtual landscape. The aggregation of the ghosts’ activities are used to draw conclusions about areas of interest, such as likely attack locations.

The DEFUSE system shows considerable promise as a “proof of concept” for anticipating potential risks in the IED domain, but needs more work. The way that the different modelling layers are integrated is an on-going research problem. The authors also wish to develop “new ways to model human behaviour that incorporate more of the non-kinetic features of human activities and decisions, for example, incorporating biased situation assessments and emotions into their behaviour”.

8 Conclusions

Adaptive algorithms, by their very nature, are ideally suited to the task of computational red teaming. Proof of concepts have been demonstrated in a number of relatively simple military scenarios. While some of this work would benefit from greater scientific rigour, some encouraging results have been achieved. The modelling environments used in these approaches, however, appear to be far short of the complexity required for IED scenarios.

At the same time, there has been considerable work, outside of red teaming per se, in developing models that exhibit a range of features that are relevant to IED scenarios. These include work on modelling the cognitive processes required for planning and analysis in asymmetric warfare, and modelling the complex social (and spatio-temporal) interactions in military risk assessment.

The challenge for the future is in combining these two. Returning once again to our criteria from Section 2, we need to develop models that have adequate descriptive richness and complexity to capture the critical interactions and decision-making that impact on IED threats. We need to identify and extract the features that are required to predict potential threats, and ascertain what kinds of adaptations of those features lead to different outcomes. And we need to examine the resulting fitness landscapes to determine what kinds of algorithms are most appropriate for zeroing in on the most dangerous threats.

References

- [1] Tucker Balch. *Behavioral Diversity in Learning Robot Teams*. PhD thesis, Georgia Institute of Technology, 1998.
- [2] Pascale Carayon and Sara Kraemer. Red team performance: Summary of findings, June 2004. University of Wisconsin Center for Quality and Productivity Improvement.
- [3] Chwee Seng Choo, Ching Lian Chua, and Su-Han Victor Tay. Automated red teaming: a proposed framework for military application. In *GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 1936–1942, New York, NY, USA, 2007. ACM.
- [4] C.L. Chua, W.C. Sim, C.S. Choo, and V. Tay. Automated red teaming: An objective-based data farming approach for red teaming. pages 1456–1462, dec. 2008.

- [5] Ibrahim Cil and Murat Mala. A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare. *Expert Syst. Appl.*, 37(2):1331–1343, 2010.
- [6] J. Crossman, R. Bechtel, H. Parunak, and S. Brueckner. Integrating dynamic social networks and spatio-temporal models for risk assessment, wargaming and planning. In *Network Science Workshop*. West Point, NY, 2009.
- [7] Brian P. Gerkey, Richard T. Vaughan, and Andrew Howard. The Player/Stage project: Tools for multi-robot and distributed sensor systems. In *ICAR2003: Proc. International Conference on Advanced Robotics*, pages 317–323, Coimbra, Portugal, 2003.
- [8] Michael Lauren and Roger Stephen. Map-aware non-uniform automata (MANA) — a new zealand approach to scenario modelling. *J. Battlefield Technology*, 5(1):27–31, 2002.
- [9] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, and Keith Sullivan. MASON: A new multiagent simulation toolkit. In *SwarmFest2004: Proceedings of the SwarmFest Workshop*. 2004.
- [10] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan. MASON: A Multiagent Simulation Environment. *SIMULATION*, 81(7):517–527, 2005.
- [11] N. Minar, R. Burkhart, C. Langton, and M. Askenazi. The Swarm simulation system: a toolkit for building multi-agent simulations. Technical Report 96-06-042, Santa Fe Institute, Santa Fe, 1996.
- [12] Michael J. North, Nicholson T. Collier, and Jerry R. Vos. Experiences creating three implementations of the repast agent modeling toolkit. *ACM Trans. Model. Comput. Simul.*, 16(1):1–25, 2006.
- [13] M Parker. What is ascape and why should you care. *Journal of Artificial Societies and Social Simulation*, 4(1), 2001.
- [14] Choo Chwee Seng, Chua Ching Lian, Low Kin Ming Spencer, and Ong Wee Sze Darren. A co-evolutionary approach for military operational analysis. In *GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, pages 67–74, New York, NY, USA, 2009. ACM.
- [15] Michael J. Skroch. Modeling and simulation of red teaming: Part 1: Why red team M & S? *Red Team Journal*, December 2009.
- [16] Praveen Kumar Tripathi, Sanghamitra Bandyopadhyay, and Sankar Kumar Pal. Multi-objective particle swarm optimization

- with time variant inertia and acceleration coefficients. *Information Sciences*, 177(22):5033 – 5049, 2007.
- [17] Stephen C. Upton, Sarah K. Johnson, and Mary J. McDonald. Breaking blue: Automated red teaming using evolvable simulations. In *GECCO 2004: Genetic and Evolutionary Computation Conference*, Seattle, Washington, 2004. Workshop on Military and Security Applications of Evolutionary Computation.
 - [18] Stephen C. Upton and Mary J. McDonald. Automated red teaming using evolutionary algorithms. In *WG31 — Computing Advances in Military OR*. 2003.
 - [19] Vesselin K. Vassilev, Terence C. Fogarty, and Julian F. Miller. Information characteristics and the structure of landscapes. *Evolutionary Computation*, 8(1):31 – 60, 2000.
 - [20] Javier Vázquez-Salceda, Virginia Dignum, and Frank Dignum. Organizing multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 11(3):307–360, 2005.
 - [21] Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(02):115–152, 1995.
 - [22] YongLiang Xu, Malcolm Yoke Hean Low, and Chwee Seng Choo. Enhancing automated red teaming with evolvable simulation. In *GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, pages 687–694, New York, NY, USA, 2009. ACM.
 - [23] Ang Yang, H.A. Abbass, and R. Sarker. Characterizing warfare in red teaming. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(2):268 –285, april 2006.
 - [24] Ang Yang, Hussein Abbass, and Ruhul Sarker. WISDOM-II: A network centric model for warfare. In *KES 2005: Ninth International Conference on Knowledge-based Intelligent Information and Engineering Systems*, volume LNCS 3683, pages 813–819. Springer, 2005.

Page classification: UNCLASSIFIED